Science Library

DEC 9 1948

Vol. 38, No. 6 December, 1948

Journal of Experimental Psychology

EDITED BY

FRANCIS W. IRWIN

UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CLARENCE H. GRAHAM, *Columbia University*
HARRY HELSON, *Bryn Mawr College*
DONALD B. LINDSLEY, *Northwestern University*
DONALD G. MARQUIS, *University of Michigan*
GEORGE A. MILLER, *Harvard University*
L. L. THURSTONE, *University of Chicago*
G. R. WENDT, *University of Rochester*
JOSEPH ZUBIN, *N. Y. Psychiatric Institute, Columbia University*

CONTENTS

The Acquisition of a Visual Discrimination Habit Following Response to a Single Stimulus: G. R. GRICE 633

The Projective Expression of Needs. II. The Effect of Different Intensities of the Hunger Drive on Thematic Apperception: J. W. ATKINSON AND D. C. MCCLELLAND 643

Studies in Spatial Learning. VI. Place Orientation and Direction Orientation: B. F. RITCHIE 659

An Attempt to Reconcile Unlearning and Reproductive Inhibition Explanations of Proactive Inhibition: B. R. BUGELSKI 670

An Experimental Analysis of Set in Rote Learning: The Interaction of Learning Instruction and Retention Performance: L. POSTMAN AND W. O. JENKINS 683

Intensity of the Conditioned Stimulus and Strength of Conditioning: I. The Conditioned Eyelid Response to Light: D. A. GRANT AND DOROTHY E. SCHNEIDER 690

Proactive Inhibition as an Effect of Handedness in Mirror Drawing: C. W. SIMON 697

Quantitative Relations Among Vernier, Real Depth, and Stereoscopic Depth Acuities: R. N. BERRY 708

The Effect of Differential Extinction on Spontaneous Recovery: A. M. LIBERMAN 722

Discrimination of the Duration of a Sound: F. M. HENRY 734

Responses to 'Meaningful' and 'Meaningless' Sounds: R. C. DAVIS 744

A Comparison of Two Indices of Palmar Sweating: M. A. WENGER AND J. C. GILCHRIST .. 757

Studies in Space Orientation. IV. Further Experiments on Perception of the Upright with Displaced Visual Fields: H. A. WITKIN AND S. E. ASCH 762

Studies of Tracking Behavior. II. The Acceleration Pattern of Quick Manual Corrective Responses: F. V. TAYLOR AND H. P. BIRMINGHAM 783

PUBLISHED BI-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

PRINCE AND LEMON STS., LANCASTER, PA.

AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, Section 538, P. L. & R., authorized September 11, 1947.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published bi-monthly by the American Psychological Association, Inc. The annual subscription rate is $7.00 (foreign, $7.25), single copies, $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1515 Massachusetts Ave., N.W., Washington 5, D. C.

This JOURNAL publishes articles which are intended to contribute toward the development of psychology as an experimental science. Experimental work with normal human subjects is favored over work with abnormal or animal subjects. Articles of a length exceeding 25 printed pages cannot be accepted.

With rare exceptions, articles are published in the order of their receipt. Authors are supplied with 50 free offprints, without covers, in the case of articles of seven or more printed pages, and with five free copies of the issue in the case of shorter contributions. One-half of the cost of tables, cuts, formulas, and author's alterations in galley proof is charged to the author.

Immediate publication can be given to articles whose authors assume the full cost of publication.

Address articles submitted for publication to the editor, Francis W. Irwin, Department of Psychology, University of Pennsylvania, Philadelphia 4, Pa.

This JOURNAL adheres to the conventions concerning bibliographies, citations, preparation of tables and cuts, etc., as described by J. E. ANDERSON and W. L. VALENTINE, The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, **41**, 345–376. Special attention is called to the style of numbers and abbreviations (p. 353). The term 'experimenter,' with its various forms, singular, plural, and possessive singular and plural, is abbreviated *E*, *E*s, *E*'s, and *E*s'; analogous abbreviations are used for the terms 'subject' (*S*) and 'observer' (*O*). All manuscripts should be typed double-spaced throughout, *including tables, footnotes, and bibliography.* Tables, figures, and footnotes should be placed on separate pages at the end of the manuscript.

Journal of
Experimental Psychology

VOL. 38, NO. 6 DECEMBER, 1948

THE ACQUISITION OF A VISUAL DISCRIMINATION HABIT FOLLOWING RESPONSE TO A SINGLE STIMULUS

BY G. ROBERT GRICE *

From the Psychological Laboratory, Brown University

INTRODUCTION

A discussion has recently arisen concerning the validity of the concept of primary stimulus generalization (**12, 6**). Lashley and Wade (**12**) in attacking the concept have based a portion of their argument upon a series of visual discrimination learning experiments in which the jumping stand technique was employed. Data from two of the experiments were presented, one of the experiments dealing with the generalization of the effects of reinforcement, and the other dealing with generalization of the effects of extinction. The present paper deals only with the experiment concerning the effects of reinforcement.

Lashley and Wade reasoned that a good test of the generalization hypothesis would be to train a group of subjects to react to a single stimulus, and then to pair that stimulus with another on the same stimulus dimension in a discrimination learning experiment. The rates of learning would then be compared for two groups, the previously reinforced stimulus now being correct for one group and incorrect for the other. According to the generalization hypothesis, the response tendency to the stimulus originally reinforced should be greater then to other stimuli on the same dimension. Thus the group now trained to choose the same stimulus should learn the discrimination at a faster rate than the group for which the previously reinforced stimulus is now incorrect. Specifically, one group of rats (the ++ Group) was trained on the jumping stand to 200 errorless trials to an 8-cm. white circle as opposed to a plain black card. A second group (the +− Group) was trained to 200 errorless trials to a 5-cm. white circle. Both groups were then trained to select the 8-cm. circle as opposed to the 5-cm. one in a discrimination problem. There were four rats in each group. In

* Now at the University of Illinois.

this test the animals with the original training to the now negative 5-cm. circle (+ −) learned the problem in fewer trials and with fewer errors than the group with the training to the still positive 8-cm. circle (+ + Group). This result was opposite to that predicted on the basis of the generalization hypothesis, and was regarded by Lashley and Wade as strong evidence against it.

One possibility that should at least be considered is that this experiment is not a good test of the generalization hypothesis. If the generalization gradient, although a genuine phenomenon, should be relatively flat for the particular dimension selected for study, the difference between the response tendencies to the reinforced stimulus and those to the other stimuli on the dimension would be small. If this were true, the difference between the two groups in the above type of experiment would be very small and a much larger number of *S*s would be required to show it clearly. Differences in the opposite direction, as obtained by Lashley and Wade, would be obtained frequently through sampling error. However, as pointed out by Hull (**6**), the fact that similar results were obtained by Lashley and Wade in other experiments (the data of which were not reported) indicates that this is probably not the explanation of the data considered here.

While the above experiment is not regarded as a particularly good test of the generalization hypothesis because of the possibility of relatively flat generalization gradients, it is more clearly related to certain other problems of learning theory. In the first place, it is relevant to the question of continuous versus non-continuous theories of discrimination learning. According to a continuity theory such as the one proposed by Spence (**16**), during the period of training to the single stimulus there should be a cumulative increase in the tendency to approach that simulus. Thus at the beginning of the discrimination problem, the response tendency to the 8-cm. circle should be stronger for the + + Group and that to the 5-cm. circle for the + − Group. Unless the difference were obscured by a flat generalization gradient, the prediction would be for the group with training to the 8-cm. circle to learn faster than the group with training to the smaller circle. According to a non-continuity theory of the type favored by Lashley (**9, 11**), and Krechevsky (**8**), the preliminary training to the single stimulus would be expected to have no effect on the rate of learning the later discrimination problem. Since a difference *was* obtained between the two groups in the Lashley and Wade experiment, and this difference favored the group with reinforcement of responses to the stimulus which later became negative, the results are contrary to both the continuity and non-continuity theories as the writer understands them. This fact has been previously recognized and pointed out by Hull (**6**).

Another theoretical problem related to the above experiment is the question of whether animals learn to respond to a relation between the stimuli in a visual discrimination problem, or whether they build up, through practice, positive and negative tendencies to the specific stimuli. A relational theory has been favored by such writers as Lashley (**10, 12**), Köhler (**7**), and Gullikson and Wolfle (**4**). The non-relational theory proposed by Spence (**16, 17**) assumes the development of excitatory and inhibitory

tendencies to specific stimuli. Assumptions are also made concerning the generalization of these tendencies. The relation of this problem to the experiment under consideration is obvious. If the eventual learning of a discrimination is dependent upon the perception of a relation between the stimuli, then previous experience with one of the stimuli alone should have no effect on the later establishment of the discrimination. Such a theory would lead to the prediction that no difference would be obtained between the two groups in Lashley and Wade's experiment. According to a non-relational theory, however, response tendencies to one stimulus would be expected to develop independently of the presence or absence of the second stimulus. Thus, such a theory would lead to the prediction that the group with preliminary training to the positive stimulus should learn the discrimination more rapidly. It should be pointed out that, here again, the experimental result of Lashley and Wade was in conflict with both theories.

In view of the theoretical importance of the Lashley and Wade experiment, the study has been repeated in its essential details. However, a somewhat different experimental technique was employed. There were several reasons why this repetition was desirable. In the first place, data were reported for only four animals in each group. In the second place, there is the difficulty in explaining the result even by the opposing theories. In the third place, there is a possibility that the results obtained were artifacts introduced by complicating factors in the jumping stand technique. While the details of the experimental procedure were not reported by Lashley and Wade, they have been described elsewhere by Lashley (**10**). Probably the most serious complicating factor is the fact that incorrect responses are punished by locking the negative card and causing the animal to fall into a net. It is well known that punished responses are more quickly eliminated than responses which are simply not rewarded. Presumably, this results from a tendency to avoid the punished response. If the animals in the + − Group had an initial tendency to respond predominately to the negative stimulus (as would be predicted by the continuity theory), they would receive more such punishments early in training than the + + Group. Another factor which makes interpretation of the results difficult is the fact that the trial and error figures reported by Lashley and Wade do not include all of the individual jumps made by the animals. According to Lashley's scoring procedure (**10**), a trial is "the behavior from the time the animal is placed on the platform until he jumps to the positive stimulus and receives food." An error is "such a trial in which the animal jumped one or more times to the negative card before choosing the positive." Thus, one 'error' included one correct jump and one or more incorrect jumps. It is possible that if all individual jumps were reported, there would be no difference between the groups, or the difference could even be in the other direction. In any case, if the + − animals

made more repetitive errors *in the early trials* of the discrimination series than the ++ animals, this fact would have been obscured by the method of scoring; but it would have been very important in determining the rates of learning of the two groups.

The experiment reported here follows the design of the Lashley and Wade experiment, but a different type of discrimination apparatus was used rather than the jumping stand, and no punishment was introduced.

Experimental procedure

Subjects.—The *S*s were 20 female albino rats from the colony maintained by the Psychological Laboratory, Brown University. Their ages at the beginning of the experiment ranged from 100 to 125 days. Ten of the *S*s were experimentally naive, and 10 had been used in a T-maze experiment. These two groups were equally divided between the experimental groups, but an analysis of the data indicates no difference between their performances in the present experiment. Aside from the equal division of naive and experienced *S*s, the rats were assigned at random to the experimental groups. The rats' diet consisted of seven gm. of Purina Laboratory Chow daily, given immediately after the experimental session. Thus, the hunger drive was a function of approximately 23 hours of food deprivation.

Apparatus.—The apparatus consisted of a specially designed discrimination box. The floor plan is shown in Fig. 1A. The box consisted of a starting alley opening into a V-shaped choice

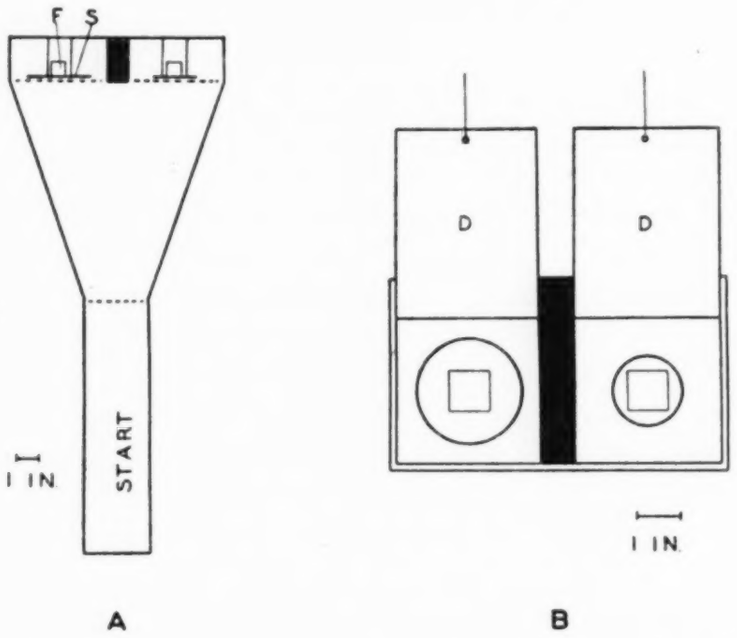


Fig. 1. Experimental apparatus. A is a floor plan of the apparatus. The dotted lines represent vertical sliding doors. The small squares (F) represent food dishes which were attached to the back of the stimulus circles (S) just below the square doors in the circles. B is a schematic cross section of the end of the apparatus containing the circles. The doors (D) could be lowered in front of the stimulus circles. The doors and the wall behind the circles were painted flat black.

chamber. The floor of the box was made of $\frac{3}{8}$ in. plywood, and the walls and doors were $\frac{1}{4}$ in. masonite. The walls were $5\frac{3}{8}$ in. high. The interior of the entire apparatus was painted a flat black. The starting alley was three in. wide and was separated from the choice chamber by a vertical sliding door. At its wide end, the choice chamber was 10 in. wide. The choice area was covered with clear glass and the starting alley was covered with hardware cloth. At the wide end of the choice chamber, opposite from the entrance, were two white circles which served as the stimuli

to be discriminated. The circles were eight and five cm. in diameter. They were cut from sheet metal, and stood out two inches in front of the rear of the box. They were 12 in. from the entrance. They were attached by brackets to vertically sliding panels which formed the rear wall of the box. The brackets were not visible from the front side of the circles. The panels and stimuli could be shifted from left to right. The centers of the circles were 5½ in. apart and 2½ in. above the floor. The circles were separated by a 1½ in. partition. Vertical sliding doors could be lowered ½ in. in front of the stimuli. At the beginning of a trial, each circle appeared in the center of a 4½ in. square, formed by the center partition, the side wall, the floor, and the lower edge of the sliding door. The top of the area behind the circles was covered with masonite, so that this area was shaded and quite dark—providing a sharp contrast to the white circles. A cross-section view of the apparatus with the stimuli is shown in Fig. 1B.

In the center of each stimulus circle was a door 1½ in. square. The door was hinged from the top at the back of the circle, so that it swung back when pushed from the front. The door fitted snugly into the circle so that when closed, the door and circle presented a flat white surface. A metal food dish, ¾ in. square and ¼ in. deep was attached to the back of the circle just below the door opening. When the door was closed, the food dish was covered by a lid, so that food could be removed only by pushing the door open from the front. The door could be opened by the *E* by means of a black thread attached to the back of the door and passing out of the apparatus through a small hole in the rear panel. The door of either circle could be locked. The back of the circles and the brackets supporting them were painted black like the rest of the apparatus interior. The inside of the food dish was painted white like the front of the circles. The illumination was supplied by a single 40-watt bulb placed 25 in. above the floor of the entrance to the choice chamber. The sliding doors were controlled by strings with counterbalance weights and were operated by the *E* who sat at the right of the table on which the apparatus rested.

Pre-training.—The rats were first trained to push open the door of the stimulus with their noses to get the reward which consisted of a .15 gm. pellet of Purina Chow. In this training the stimulus later to be used in the single stimulus series was used. On the first day of training the rat was placed in the choice compartment with the stimulus on the left and with the small door in the stimulus open. A piece of hardware cloth was placed immediately back of the stimulus to eliminate exploratory activity in the covered area behind the stimulus. The animal was left in the apparatus until it had eaten six pellets of food from the food dish, or for a minimum of 15 min. On the second day the animal was given 10 trials running to the stimulus from the starting alley. The hardware cloth behind the stimulus was still used on this day. On the first trial the food door was held open by the *E*. On successive trials, the door was closed a little more each time, so that by the tenth trial all animals had learned to push open the door and get the food when the door was completely closed. On the day following this it was occasionally necessary to hold the door open briefly by means of the thread after the animal had opened it. This was done in order to prevent the door from closing and frightening the animal before the food was obtained. After that day, however, no further aid was necessary. The stimuli on the pre-training day appeared five times on the right and five times on the left in the order LRLRLRLRLR. When the animal had obtained the food, the sliding door in front of the stimulus was closed and the animal was allowed to eat in the choice compartment. The door on the side opposite to the stimulus was always closed.

Single stimulus training.—On the day following pre-training, the training to the single stimulus began. The animals received 20 trials per day for 10 days (200 trials in all) running to a single stimulus for the food reward. Ten *S*s were trained to the 8-cm. circle (+ + Group), and 10 were trained to the 5-cm. circle (+ − Group). The stimulus appeared half of the time on the right and half on the left in the order RLRRLLRLLRLRLLRRLRRL. The side opposite to the stimulus was filled by the closed black sliding door. This is similar to the Lashley and Wade procedure of using a black card on the negative side. The procedure was to place the animal in the starting alley, raise the door, and allow the animal to go to the stimulus and obtain the food. The door was then lowered in front of the stimulus. While the animal ate, the apparatus was prepared for the next trial. When the animal had eaten, it was then replaced in the starting alley and the next trial was run immediately.

Discrimination training.—On the day following the completion of the training to single stimuli, both groups began training on the discrimination problem. The 8-cm. circle was correct and the 5-cm. circle was incorrect for both groups. In order to control any possibility of olfactory cues, food was placed in both the positive and negative food dishes, but the door of the smaller stimulus was locked. In this situation, both of the sliding black doors were open so that the stimuli appeared as in Fig. 1B. The positive stimulus appeared equally often on the left and right in the

order RLRRLLRLLRLRLLRRLRRL. If the animal's nose touched the door of the negative stimulus, both black doors were immediately lowered and the trial was scored as an error. No punishment was used, and no correction was permitted. If the animal made a correct response, both black doors were closed, and the animal was allowed to eat. There were 20 trials per day, and all animals were run until they reached a criterion of 18 out of 20 successive trials correct with the last 10 perfect.

Results and Discussion

The experimental results are summarized in Table I. In terms of both number of errors and number of trials required to reach the criterion of learning, the ++ Group showed marked superiority over the +− Group. The skewness of the distributions and the non-homogeniety of the variances make the *t*-test inappropriate for testing

TABLE I

Number of Errors Made by Each *S*, and Number of Trials Required to Reach the Criterion of Learning in the Discrimination Problem

Means and medians are also given for each group.

++ Group			+− Group		
Rat	Errors	Trials	Rat	Errors	Trials
12	13	31	11	11	36
8	14	29	3	21	40
9	13	29	15	29	99
18	10	34	21	14	48
19	12	28	13	53	135
30	18	42	1	15	49
4	8	17	2	30	73
41	24	41	44	60	126
42	10	24	45	25	72
43	7	20	46	12	38
Median	12.5	29.0	Median	23.0	60.0
Mean	12.9	29.5	Mean	27.0	71.6

the significance of these differences. Consequently, non-parametric tests of significance were applied. According to the *d*-test of Festinger (**3**) the difference between the two distributions of error scores is significant at the five percent level of confidence, and the difference between the trial scores is significant at the one percent level. According to the *u*-test of Mann and Whitney (**14**), these two differences are significant at the one and 0.1 percent levels of confidence respectively. The difference in rate of learning of the two groups is shown graphically in Fig. 2. These group learning curves indicate a marked superiority of the ++ Group over the +− Group in rate of learning to select the larger circle. It will be noted that both groups started at levels of greater than 50 percent choice of the larger stimulus. This was apparently the result of a size preference which is typical in size discrimination problems.

Clearly, these results are contrary to those of Lashley and Wade. Undoubtedly, this difference is the result of the differing experimental procedures used. As previously mentioned, one difference was the differing scoring procedures. However, it is not possible to evaluate the importance of this difference since Lashley and Wade did not report the number of individual jumps on their 'error' trials. The most important difference is probably the absence of punishment for errors in the present experiment as contrasted with its presence in the jumping stand procedure. One possible explanation of the differing results is that the animals in Lashley and Wade's + − Group quickly acquired an avoidance tendency to the 5-cm. circle as a result of

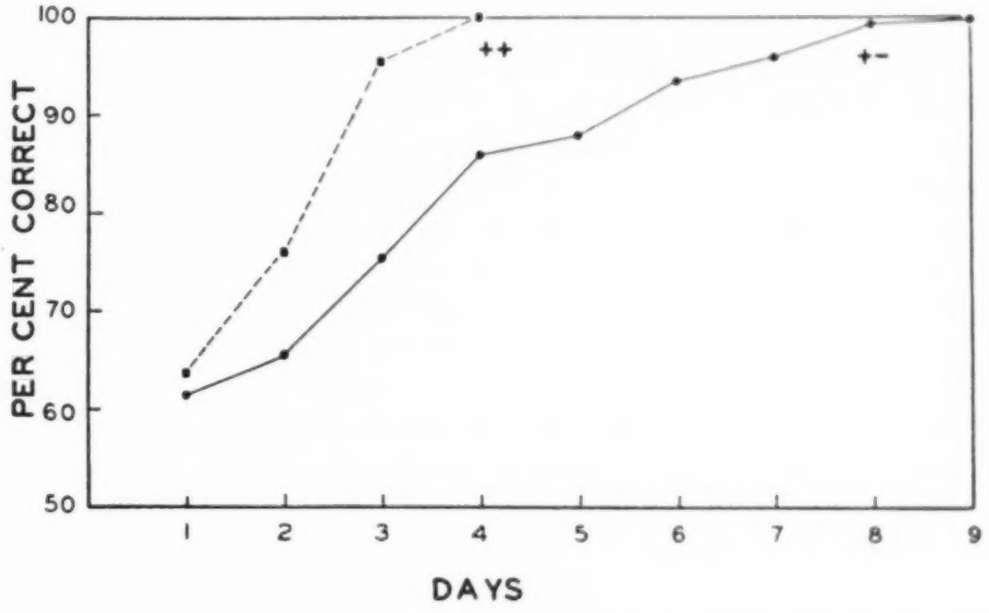


FIG. 2. Group learning curves on the discrimination problem for the two experimental groups. The percent of choices of the larger circle is shown for each day's session.

punished responses. As previously suggested, the + − Group would be expected to acquire an avoidance tendency more rapidly than the ++ Group because it would be expected to make more early jumps to the negative stimulus. Assuming that, due to generalization, the positive excitatory tendencies to the larger circle did not greatly differ for the two groups, this greater tendency to avoid the negative stimulus could easily account for the more rapid learning of the + − Group. Whether or not this analysis of the conflicting experimental results is the correct one, since this experiment purports to deal with generalization of the effects of reinforcement, it would appear much better to study the phenomenon in a situation uncomplicated by the effects of punishment. In the present experiment no punishment was involved. Incorrect responses were merely not reinforced. Such a procedure is less likely to lead to the rapid development of avoidance

reactions. Any difference in positive excitatory tendencies to the two stimuli which may have developed during the training to a single stimulus would thus be more likely to show up in the acquisition of the discrimination habit. On the basis of Lashley and Wade's prediction from the theory of generalization, the ++ Group should have a greater reaction tendency to the 8-cm. circle and the +− Group should have a greater reaction tendency to the 5-cm. circle, since their initial training was to these stimuli respectively. On the basis of this assumed difference, the ++ Group would be expected to learn more rapidly than the +− Group to select the larger of the two circles. This is exactly the result which was obtained in the present experiment at a high level of statistical significance. Thus, the experiment verifies the prediction made by Lashley and Wade on the basis of the generalization theory which they rejected. Since a difference between the groups was obtained, it is indicated that the generalization gradient for the size dimension is not flat, and that greater habit strength is attached to the specifically reinforced stimulus. Since the present experiment involved more *S*s, produced statistically significant differences, and was uncomplicated by the effects of punishment, in the opinion of the writer, it provides a more adequate test of the theory than the earlier experiment of Lashley and Wade.

This experiment is even more directly applicable to the problem of continuity versus non-continuity theories of descrimination learning. According to the non-continuity view, supported principally by Krechevsky (**8**) and Lashley (**11**), learning in a descrimination problem does not occur until the animal finally selects out the appropriate aspects of the stimuli and responds to them. Krechevsky describes this as adopting the correct hypothesis. Practice prior to this is assumed to have no effect upon the final learning. Since in training to a single stimulus there would be no possibility of adopting an hypothesis applicable to the discrimination problem, and since learning is not considered to be cumulative, the prediction of such a theory would be that there would be no difference between the ++ and the +− groups in the present experiment. However, as previously mentioned, the continuity theory assumes that excitatory tendency or habit strength builds up gradually with each reinforced response to a stimulus. Rather conclusive evidence has already been presented to support this theory (**1, 2, 13, 15, 18, 19, 20**). On the basis of the training to the single stimuli in this study, this theory would lead to the prediction that the ++ Group would learn the discrimination problem more rapidly than the +− Group, since the training of the ++ Group was to the positive stimulus. The results of the experiment, then, are further confirming evidence in support of the continuity theory as opposed to the non-continuity theory.

This experiment also tends to support the view that, in size discrimination at least, learning is dependent on the establishment of response tendencies to specific stimuli rather than to a relationship between them. Since training to single stimuli alone did significantly affect the later acquisition of the discrimination habit, it is obvious that reactions to specific stimuli are involved in forming the discrimination. No opportunity for learning anything about the size relation between the stimuli was possible during the single stimulus training, yet the training had a differential effect upon the two groups. Training to the specific stimulus which later bacame positive led to more rapid learning than training to the specific stimulus which later became negative. This is much more in accord with the interpretation in terms of learning responses to specific stimuli rather than to relationships between them. There is nothing in the present experiment, however, to indicate that the learning of relationships between stimuli is impossible.

In summary, it might be said that the results of this experiment are in accord with the stimulus-response-reinforcement type of learning theory supported by Hull (**5**) and Spence (**16, 17**). Such a theory assumes that learning consists of the gradual acquisition of habit strength between stimuli and responses as the result of successive reinforcements. This habit strength may generalize to stimuli similar to the one with which the original reinforcement occurred. However, the greatest amount of habit strength is assumed to be associated with the specifically reinforced stimulus.

Summary

1. In a specially designed visual discrimination apparatus, 10 white rats each made 200 rewarded responses to a single white circle eight cm. in diameter. Another group of 10 rats made 200 rewarded responses to a five centimeter white circle. Both groups were then trained in a size discrimination problem in which the 8-cm. circle was positive and the 5-cm. circle was negative.

2. The group with the preliminary training to the 8-cm. circle learned the discrimination problem at a significantly faster rate than the group with the preliminary training to the 5-cm. circle.

3. The data are interpreted as being in accord with the theory of generalization.

4. The data also are interpreted as confirming evidence for the continuity theory as opposed to the non-continuity interpretation of discrimination learning.

5. The result further supports the view that size discrimination learning consists of the establishment of reaction tendencies to the specific stimuli rather than to a relationship between them.

(Manuscript received December 22, 1947)

References

1. Bollinger, P. An experimental study of the pre-solution period of discrimination learning in the white rat. M. A. thesis deposited in the State University of Iowa Library, 1940.
2. Ehrenfreund, D. An experimental test of the continuity theory of discrimination learning with pattern vision. *Amer. Psychologist*, 1947, **2**, 316. (Abstract)
3. Festinger, L. The significance of difference between means without reference to the frequency distribution function. *Psychometrika*, 1946, **11**, 97–105.
4. Gulliksen, H., & Wolfle, D. L. Theory of learning and transfer. *Psychometrika*, 1938, 3, 127–149, 225–251.
5. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
6. Hull, C. L. The problem of primary stimulus generalization. *Psychol. Rev.*, 1947, **54**, 120–134.
7. Köhler, W. Aus der Anthropoidenstation auf Teneriffa. IV. Nachweis einfacher Structurfunktionen beim Schipansen und beim Haushuhn: Über eine neue Methode zur Untersuchung des bunten Farbensystems. *Abh. preuss. Akad. Wiss.*, 1918, Berlin, 1–101.
8. Krechevsky, I. 'Hypotheses' in rats. *Psychol. Rev.* 1932, **39**, 516–532.
9. Lashley, K. S. *Brain mechanisms and intelligence.* Chicago: Univ. Chicago Press, 1929.
10. Lashley, K. S. The mechanism of vision. XV. Preliminary studies of the rat's capacity for detail vision. *J. gen. Psychol.*, 1938, **18**, 123–193.
11. Lashley, K. S. An examination of the 'continuity' theory as applied to discrimination learning. *J. gen. Psychol.*, 1942, **26**, 241–265.
12. Lashley, K. S., & Wade, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, **53**, 72–87.
13. Leeper, R. The continuity, non-continuity issue in learning—an appraisal and a re-formulation. *Amer. Psychologist*, 1947, **2**, 316–317. (Abstract)
14. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Annals math. Statistics*, 1947, **18**, 50–60.
15. McCulloch, T. L., & Pratt, J. C. A study of the pre-solution period in weight discrimination by white rats. *J. comp. Psychol.*, 1934, **18**, 271–290.
16. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, 43, 427–449.
17. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430–444.
18. Spence, K. W. Analysis of the formation of visual discrimination habits in chimpanzee. *J. comp. Psychol.*, 1937, **23**, 77–100.
19. Spence, K. W. Continuous versus non-continuous interpretations of discrimination learning. *Psychol. Rev.*, 1940, **47**, 271–288.
20. Spence, K. W. An experimental test of the continuity and non-continuity theories of discrimination learning. *J. exp. Psychol.*, 1945, **35**, 253–266.

THE PROJECTIVE EXPRESSION OF NEEDS. II. THE EFFECT OF DIFFERENT INTENSITIES OF THE HUNGER DRIVE ON THEMATIC APPERCEPTION[1]

BY JOHN W. ATKINSON AND DAVID C. McCLELLAND

Wesleyan University

The first experiment in this series (7) attempted to measure the effects of different intensities of the hunger drive on perception for which the objective determinants had been reduced to a minumum. Its purpose was to seek any principles which might govern the relationship between need intensity and its expression in perception. The results provided clues as to how perceptual material should be interpreted to diagnose the strength of a drive such as hunger.

But most so-called projective methods do not deal with perceptual material unless, like the Rorschach, they are concerned more with the formal than the contentual aspects of phantasy. Therefore, the next step is to see whether the findings with perceptual material apply to a prejective method more nearly like those ordinarily employed—the Thematic Apperception Test, for example. Several previous attempts have been made to measure the effect on imaginative productions of specific experimentally induced conditions (**2, 5, 11**), but none has dealt with the problem very extensively or attempted to control the amount of the variable, the effect of which is to be measured.

In the present experiment the intensity of the hunger drive was controlled as in the previous experiment by depriving the subjects of food for one hour, four hours, and 16 hours. They then wrote out short stories about pictures drawn mostly from Murray's Thematic Apperception Test (referred to hereafter as the TAT). This design permitted the direct comparison of contentual aspects of the stories told under the influence of three degrees of intensity of the hunger drive.

Procedure

The *Ss* were drawn from Naval personnel attending the submarine training school at the U. S. Navy Submarine Base at New London, Conn., obtained as previously described (7). Food deprivation was also controlled as before. In fact, some of the *Ss* used in the previous experiment participated in this experiment after concluding the tests of perception. In all there were 44,

[1] This project was made possible by a grant from the Office of Naval Research. The authors are grateful for this aid and particularly for the cooperation of Captain C. W. Schilling, formerly director of the Medical Research Laboratory, U. S. Naval Submarine Base, New London, Conn., and his staff.

22, and 37 *Ss* in the one, four, and 16 hour deprivation groups respectively who took the entire test and whose records were analyzed. 30 more *Ss* had to be discarded because they did not meet the experimental hunger conditions or because their stories were too incomplete to work with. Of the 103 *Ss* whose stories were analyzed, 22 had to be discarded for most of the results reported in the paper because they did not see two of the pictures.

The pictures used are described in Table I with the code letter used to identify them. They

TABLE I

DESCRIPTION OF PICTURES, GIVING CODE LETTER AND SOURCE

A. Four 'bums' resting (TAT 9 BM in 1943 edition)
B. 'Beachcomber' (TAT earlier edition, cf. 10, p. 542)
C. 'Father' talking to 'son' (TAT 7 BM)
D. Woman looking into room (TAT 5)
E. Restaurant balcony scene—man and woman looking out window, table in foreground (Meier-Seashore Art Judgment Test picture No. 84)
F. Older man in black hat, younger blond man, piece of raw meat with knife beside it (especially designed for this experiment—three picture elements in color not joined together in a single scene)
G. Man standing under lamp (TAT 20)
I. Men in tavern or coffee shop talking (especially designed for this experiment)

were chosen to suggest all types of situations relating to hunger—satiation (A), deprivation (B), invitation to eat (D), place of eating (E), food (F) etc. Slides were made of all these pictures (except F) and were visibly projected under conditions described elsewhere (7). F was projected by reflection. There were two orders of presentation as follows: pictures A B C D E F G I and D A F B E G I (picture C was dropped as it gave no food responses). As stated above, 22 subjects did not see pictures G and I, but their results are included in the analysis of characteristics of various pictures. Most of the *Ss* saw the pictures in the second order but 13 in the one hour group and 15 in the 16 hour group saw them in the first order.

The procedure is best described from the instructions given the subjects after they had entered the room in groups of 7–15 and taken seats facing the screen.

> "Please fill out the information blank on the top of your answer sheet [which asked simply for name, rank, dates of birth and induction].
>
> "This is a test of your creative imagination. Six pictures will be projected on the screen. You will have twenty seconds to look at the picture and then five minutes to write what you think about it. Notice that on each of the six pages on your answer sheet corresponding to the six pictures to be shown, the same four questions are asked. Try to spend about one minute answering each question. I'll help you keep track of the time so that you can answer all four questions for each picture. I'll tell you when to go on to the next question. You'll have time to go back and finish up before the next picture is shown.
>
> "Obviously there are no right or wrong answers, so you may feel free to make up any kind of a story about the picture that you choose. The more vivid and dramatic the better. I would suggest writing as fast but as legibly as you can, in order to make your story about the picture as vivid and detailed and imaginative as you can. Remember this is a test of creative imagination. Do not merely describe the picture. Anyone can do that. Make up a story about it. The four questions will guide your thinking so that you may cover all the ground in the time required.
>
> "You may start writing as soon as you like after the picture is projected. The lights will be turned on after twenty seconds. Don't wait more than 15 seconds to start or you won't have time to do a complete job. The four questions are:
>
> 1. What is happening? Who are the persons?
> 2. What has led up to this situation? That is, what has happened in the past?
> 3. What is being thought? What is wanted? By whom?
> 4. What will happen? What will be done?
>
> "These questions cover the whole plot of the story pretty well. If you need more room for any question use the back of the paper. Make the stories as interesting as you can. Work as fast as you can—feel free to make up any kind of story at all. Be as imaginative as

your are able to be. Does every one understand the directions? No questions will be answered after we once start. Also please do not say anything when the pictures are shown."

Special instructions were necessary to introduce picture F which was made up of three disconnected parts. The following comment was therefore made just before it was shown in the third or sixth position.

"This picture is a little different. There will be a number of disconnected pictures, persons or objects that you can fit into a story. On the basis of what you see in this next picture, work out a story including what you see and anything else you might want to bring into it. Be imaginative."

In choosing the materials and in following the procedure just outlined every attempt was made to follow previous practices so far as possible in order to eliminate results peculiar to a different technique. For instance, the four questions spaced apart on the record sheets are those ordinarily raised at the beginning of a TAT session (**10**, p. 532). They were repeated here to remind the *S*s of the various aspects of a story which were wanted and to facilitate the scoring of the records.

Scoring the records proved to be the major difficulty in obtaining results, as seems always to be the case with the experimental use of projective methods. There are any number of solutions to the difficulty which have been proposed (**4**, **9**, **12**), but there is no standard method of scoring which could be adopted. This is particularly unfortunate because the results obtained by one method of scoring cannot be compared with those from any other method. Since there was no adequate methodological basis for choosing a system, one was finally worked out for use here which was based in part on Murray's need analysis, in part on the customary analysis of a behavior sequence (need, instrumental act, goal response), and in part on results obtained in the first experiment in the series.

More than 50 stories were read and considered before any catagory of analysis was formulated. The effort was to construct categories which could be identified by the *S*'s actual words with a minimum of inference on the part of the judge. The categories finally used follow with a brief description of each:

F I *Food related imagery.* Any reference to anything having to do with food or eating. This is the criterion for further analysis. It indicates that food is at least incidental to the story being told. All stories not checked in this category were not analyzed any further.

F th *Food thema.* Food getting or food enjoying activity is the central plot of the story. Someone goes to the store to buy meat for supper. A family is getting ready to eat, etc.

D th *Food deprivation thema.* Deprivation of food is the central plot of the story. The deprivation story dwells on the threatening situation, the shortage (the black market), the down and out person on the verge of starvation etc.[2] Instrumental acts are aimed at removing the deprivation. In food stories there may be difficulties in getting food but unless they constitute the primary theme or emphasis of the story, they were not scored D th (but were scored D).

O th *Other themas.* All themas not classified as F or D, i.e. food was not central to their plot.

N *Need food.* Someone in the story wants food (as actually stated, not inferred). This category includes most food themas and also all cases where someone wants food but the central thema of the story is about something else. The need must be stated and is not inferred from instrumental food-getting activity.

D *Food deprivation.* Any shortage, scarcity, blocking by external agent etc. leading to food deprivation, temporary or permanent. This includes all Deprivation Thema stories and also all other cases of food deprivation whether food is central to the plot or not.

I, f or d, +, −, or o. *Instrumental activity* aimed at getting food (I f) or at removing deprivation (I d) which is successful (+), unsuccessful (−) or of unknown outcome (o). Someone in the story must do something either to get food (I f) or to overcome the deprivation such as the black market (I d).

[2] There were scoring difficulties with picture B because many *S*s wrote severe deprivation stories for it (shipwreck, airplane crash etc.) but it was hard to decide whether food deprivation was primary. A general rule of thumb was finally established that it would be scored D th if only *one* other need beside food was mentioned (e.g., water or rest). If two or more other needs were mentioned it was scored D but not D th.

G *Goal activity* (terminal behavior). Someone is eating, will eat immediately after the action or has just finished eating. Mealtime words such as lunch and dinner were scored G.

H, i, s or p. *Hostility*, either instrumental (i), subjective (s), or pressive (p). In the first instance, aggression (stealing, violent action) is used as the instrumental activity to get food or overcome the deprivation. In the second, one or more characters in the stories feels angry or aggressive, but doesn't do anything about it. In the third, there is a hostile or aggressive press (either actual or imagined), usually the source of the deprivation. If someone stole food (H i) and was met with defensive hostility, this was not scored H p. This category contains only easily identifiable hostility and aggression.

A, s or g. *Anxiety* over survival (s) (fear of death or starving) or over guilt (g) as when food has been stolen, black market purchases made, etc.

P *Phantasy*. Someone wishing for or dreaming about getting food. This does not include the ordinary thoughts or intents that precede action.

WF *Wish fulfillment*. The person in need of food is satisfied by chance, an act of fate, a handout, etc. but not by his own efforts.

S *Substitution*. Some other need is satisfied at the end of the story in place of the blocked or partly blocked need for food. This does not include satisfactions from overcoming food deprivation.

F p *Friendly press*. Someone in the story is invited to eat or helped to get food.

This system of categories was designed to cover the major aspects of a behavior sequence involving food-to wit, the motivating state, the kind of situation (deprivation or food-seeking), the instrumental activities (active or passive wish fulfillment, real or unreal, hostile or problem solving) and the goal response activity (amount of eating and enjoyment of food). Its utility can perhaps best be illustrated by its application to three typical sample stories. All of the following occurred in response to picture F in the 16 hour deprivation group.

Story 1

"The persons are a man from the black market and an honest citizen of a small Southern town. The citizen hasn't had any meat for a couple of weeks and the man from the black market knows it. The man from the black market thinks he can sell the meat for twice what it is worth. The citizen doesn't want to buy from the black market but he is wanting the meat. The citizen doesn't buy the meat and reports it to the police."

There is food related imagery (F I). The central thema is clearly the need of the 'honest citizen' for the meat (F th) which he wants (N). Deprivation (D) is stated in the words 'hasn't had any meat for a couple of weeks.' The instrumental acts do not get food (I f−) but do get rid of at least part of the reason for the shortage—the black marketer (I d+). There is also some substitute satisfaction (S) implied for doing the right thing ('doesn't want to buy. . .reports it to the police').

Story 2

"The man in black is a detective and he is studying the meat to see if he can get a clue to the black marketer who is the man in grey. There has been a shortage of meat and also a large cost when available. The detective thinks he can catch the crook—he wants a clue. The detective will investigate anything that arouses his suspicion and he will catch the crook."

This story has the necessary food related imagery (F) but it is clearly about the shortage of meat and is therefore a deprivation story (D th). No food need is stated and the action is simple. He catches the crook (I d+).

Many of the stories were more complicated than these which have been chosen to illustrate some of the essential differences in the classification categories. Story 3 illustrates some of the difficulties in scoring that occurred.

Story 3

"The man on the top was a preacher in a Catholic Church and the younger man a member of his parish. It being Friday, all Catholics should sustain (sic) from eating any flesh of any animal. Upon visiting the member's home and discovering that they are having meat to eat he is shocked. The pastor will tell the younger man his fault, and how he is not a good Catholic. The man will outwardly say he's sorry, but he will do it again because he wants what he wants when he wants it."

This is a food thema story which involves actual eating (G) as well as the need for food (N). There is interference on the part of the priest (D) but the response is successful in overcoming it (I d+). The problem here lies partly in deciding that this is not centrally concerned with a threatened deprivation (a decision based mostly on the actual eating as compared with absence of real deprivation) and partly in deciding that the priest's intervention is deprivation and not hostile press (a decision based on the absence of actual aggression). Most calssification problems were of this sort and usually they could be resolved on some such basis as these.

To arrive at the final scores, one *E* (JWA) first classified all the stories. The other *E* (DCM) then classified independently all the stories in response to picture F (which accounted for nearly half the food stories given). A running tally showed there was agreement on about 75–80 percent of the classifications. However, most of the disagreements came in the hostility category in trying to decide whether it was i, s, or p. In the main categories (F I, N, G, etc.) there was disagreement in less than five percent of the cases. Then the two judges decided together which of two classifications a response should receive. During discussion the hazier definitions were clarified so that there was ready agreement on practically all calssifications, once the definitions were fully understood by both judges in the same way. Finally the same procedure was adopted with the remainder of the stories. In general only the most obvious points were scored—that is, there had to be actual words that supported each tally. The scores should therefore be considered as conservative estimates of the most obvious trends. The judges knew to which group the *Ss* belonged, but could not have easily influenced the results as they had no clear-cut expectations as to how the results would come out.

Results

The major results of the experiment are presented in Tables II (for the situation) and III (for the reactions) in terms of the percentage of the subjects in each deprivation group showing a given characteristic at least once in the seven stories told. This method of tabulating the data is somewhat insensitive, as it does not count the times beyond the first that a characteristic appears. Other methods were tried, such as the number of times a characteristic appeared out of the total possible number of times, but they yielded such small percentages that, although the results were the same, the major trends in the data were more difficult to see. Furthermore the differences among

TABLE II

Number and Ppercentage of *Ss* Showing Various Kinds of Situational Story Characteristics at Least Once in the Total Test of Seven Pictures

Chi-square = 3.8 and 6.6 at the 5 percent and 1 percent levels of significance.

Characteristic	1 Hour Deprivation $N = 38$		4 Hour Deprivation $N = 21$		16 Hour Deprivation $N = 22$		Chi-square of Differences		
	N	%	*N*	%	*N*	%	1–4 hr.	4–16 hr.	1–16 hr.
Food imagery	35	92.1	18	85.7	20	90.9			
Food central	25	65.8	13	61.9	20	90.9		3.52*	4.69
F thema	19	50.0	13	61.9	13	59.1			
D thema	7	18.4	3	14.3	11	50.0		6.24	6.62
Block**	2	5.3	1	4.8	2	9.1			
Need food	21	55.3	18	85.7	18	81.8	5.60		4.32
F deprivation	19	50.0	11	52.4	18	81.8		4.24	5.97

* Yates's correction used.

** Blocks do not include *Ss* who blocked on more than half the picture and hence were voided.

the groups obtained by the method used in Tables II and III, when compared with mean scores wherever there were enough cases to calculate a mean, were the same as obtained by the use of mean scores when the total number of appearances of a characteristic were taken into account.[3]

Table II, which summarizes the data on the statement of the situation, shows that there is no overall increase in food imagery as hunger increases. Since this is a case where the percentages are high and therefore possibly insensitive to small differences, the mean num-

TABLE III

NUMBER AND PERCENTAGE OF *Ss* SHOWING VARIOUS KINDS OF REACTIONS IN THEIR STORIES AT LEAST ONCE IN THE TOTAL TEST OF SEVEN PICTURES

Characteristic	1 Hour Deprivation $N = 38$		4 Hour Deprivation $N = 21$		16 Hour Deprivation $N = 22$		Chi-square of Differences		
	N	%	*N*	%	*N*	%	1–4 hr.	4–16 hr.	1–16 hr.
Instrumental Activity	23	60.5	14	66.7	20	90.9		2.83*	6.33
f +	12	31.6	10	47.6	9	40.9	1.49		
d +	5	13.2	2	9.5	12	54.6		10.89	11.75
f− or d−	5	13.2	5	23.8	8	36.4			3.08
fo or do	7	18.4	4	19.1	4	18.2			
Goal activity	28	73.7	13	61.9	8	36.4	.89	2.81	8.09
Hostility	17	44.7	12	57.1	11	50.0			
Instr.	8	21.1	6	28.6	7	31.8			
Subj.	5	13.2	4	19.1	2	9.1			
Press	6	15.8	5	23.8	3	13.6			
Friendly Press	9	23.7	7	33.3	0	0.0		6.56*	4.39*
Wish phantasy	11	29.0	4	19.1	4	18.2			
Wish fulfillment	8	21.1	4	19.1	4	18.2			
Anxiety Guilt	8	21.1	3	14.3	4	18.2			
Survival	5	13.2	3	14.3	7	31.8			1.99*
Substitution	2	5.3	3	14.3	2	9.1			

Chi-square = 3.8 and 6.6 at 5 percent and 1 percent levels of significance.
* Yates's correction used.

ber of times food imagery appeared in the seven pictures for each of the three groups was also calculated. The means for the one hour, four hour and 16 hour groups were respectively: 2.32, $\sigma_m = .20$; 2.29, $\sigma_m = .28$; and 2.23, $\sigma_m = .28$. The differences are slight and insignificant. When, however, the relevance of food to the central plot of

[3] A rough check on the stability of these percentages was made by dividing the 38 *Ss* in the 1 hour group into two groups of 19 *Ss* and correlating the frequencies obtained by the two groups for each of 21 categories. The product moment correlation was .89 (uncorrected) which indicates a fair degree of reliability of the data as far as results on groups of around 20 *Ss* are concerned. This also provides some check on the ability of the judges to classify the characteristics consistently.

the story is considered, a very different picture appears. Here there is a decided increase in the number of times the plot revolves around food as hunger increases. Further analysis[4] reveals that the increase lies in the food-deprivation plots and not in the food getting plots. This last point is confirmed by the increase in food deprivation whether central or not which appears in the last line of the table. In general, these differences appear only between the one and four hour groups and the 16 hour group. The only difference of any significance between the one hour and the four hour groups was in the number of times the need for food was expressed.

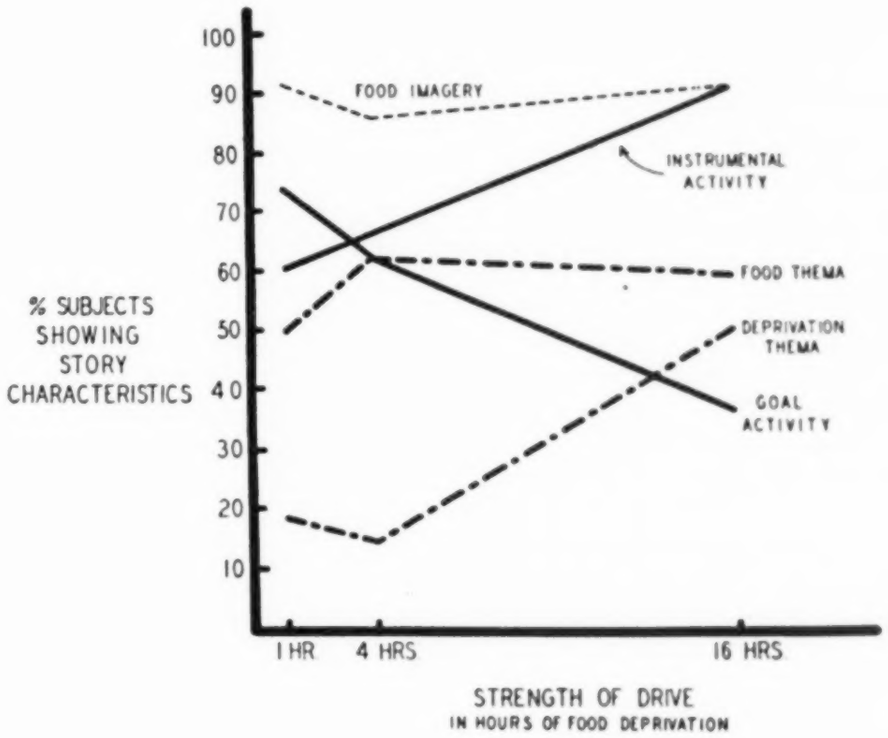


FIG. 1. Percentage of *Ss* showing selected food related story characteristics as a function of increasing hunger

Table III continues the comparison for some of the reactions of the characters in the stories to a behavior sequence involving food. The differences which are significant or fairly large are checked by means of chi-square. As hunger increases, there is an increase in instrumental activity which is focussed not on food getting but on removing the source of deprivation of food. On the other hand, there is a marked decrease in goal activity (eating) and in press actively promoting goal activity (invitations to eating). None of the other differences is re-

[4] Note that the sum of the percentages of the two sub-categories is usually more than the percentage in the overall category. This is because the overall category includes those *Ss* who show *either* F th *or* D th, while some may show *both*, a fact which appears separately in the sub-classifications.

liable, though several are in an understandable direction (e.g., increase in anxiety over survival, increase in instrumental hostility as a means of getting food, decrease in phantasies just wishing for food, etc.).

Some of the major results are presented graphically in Fig. 1, which shows, first, the failure of food imagery to increase; second, the complementary shift from goal to instrumental activity; and third, the flatness of the food getting thema curve as compared with the sharp increase between four and 16 hours deprivation for the deprivation thema curve which is characteristic of all deprivation related measures (D th, D, I d+, f− or d−).

On the basis of these shifts a scoring method was devised for summarizing each person's standing in relation to his projected need for food. The following categories were scored +1: D th, N, D, and I f− or d−. I d+ was scored +2 (because of the size of its chi-square) and G and F p were scored −1. A characteristic was scored each time it appeared in the total test. The resulting total will be referred to as the *Need Food score* (NF Score). Table IV summarizes

TABLE IV

MEAN NEED FOOD (NF) SCORES OF THE THREE DEPRIVATION GROUPS WITH ESTIMATED RELIABILITIES OF THE DIFFERENCES AMONG THEM

Hours Food Deprivation	Mean NF Score	σ_m	Difference and Probability of its Occurring by Chance with			
			4 hr.	*P*	16 hr.	*P*
1 hour	.74	.36	.83	<.10	3.31	<.01
4 hour	1.57	.52	—	—	2.48	<.01
16 hour	4.05	.66	—	—	—	—

the mean NF scores for the three deprivation groups. As far as means are concerned, the one and four hour groups are significantly lower on the NF scale than the 16 hour group. A more interesting question is whether the group from which the *S*s came could be correctly identified from the NF score alone. Obviously the four hour group overlaps so much with both other groups that the members of this group cannot be very successfully identified from their NF scores. However if the 60 subjects in the one and 16 hour groups were thrown together and sorted out with those scoring one or less in the one hour group, those scoring two in neither, and those scoring three or more in the 16 hour group, 65 percent of the *S*s would be correctly placed on the basis of their projections, 18 percent would be unclassified, and 17 percent would be wrongly classified ('false positives or negatives'). Of course, it is possible to eliminate the wrong classifications by changing the criteria, but only by greatly increasing the number un-

classified. For example, none of the *Ss* in the 16 hour group had a minus NF score while nine out of the 38, or 24 percent of the one hour group had minus scores. It can therefore be said with certainty (within the limits of this sample) that anyone receiving a minus NF score is not in a state of hunger, i.e., he has not been deprived of food for as long as 16 hours. On the other hand, it can be said almost with as much certainty (there is one exception) that anyone scoring five or more (as 10, or 46 percent, of the 16 hour group did) has not eaten recently (within the last one to two hours).

TABLE V

PERCENTAGE OF ALL THE *Ss* IN THE ONE HOUR DEPRIVATION GROUP ($N=44$) AND IN THE 16 HOUR GROUP ($N=37$) SHOWING SELECTED CHARACTERISTICS FOR DIFFERENT PICTURES

	F I	N	D	I	I d	G	Fp	Mean NH Score
Picture A (Four 'bums' resting)								
1 hour	50.0	6.8	2.3	9.1	2.3	27.3	0.0	−.14
16 hour	35.1	10.8	8.1	5.4	0.0	18.9	0.0	.03
Picture B ('Beachcomber')								
1 hour	27.3	15.0	18.2	11.4	5.4	4.6	0.0	.45
16 hour	32.4	21.6	32.4	24.3	16.2	2.7	0.0	1.03
Picture D (Woman looking into room)								
1 hour	13.6	2.3	0.0	2.3	0.0	6.8	0.0	−.11
16 hour	18.9	5.4	0.0	10.8	0.0	8.1	10.8	−.16
Picture E (Restaurant balcony scene)								
1 hour	22.7	2.3	0.0	2.3	0.0	18.2	2.3	−.21
16 hour	13.5	2.7	0.0	5.4	0.0	10.8	0.0	−.08
Picture F (Two men and piece of meat)								
1 hour	81.8	36.4	27.3	43.2	2.3	27.3	18.2	.39
16 hour	83.8	54.1	40.5	64.9	21.6	10.8	2.7	1.62
Chi-square		2.6		3.8	5.8*	3.4	3.4*	

* Yates's correction used.

Table V presents a breakdown of scores for some of the main categories on five of the pictures for which data were available on the larger number of *Ss*.[5] Also for the sake of simplicity only the percentages of *Ss* showing a characteristic in the one hour and 16 hour groups are presented. Several facts immediately emerge from these data. In the first place, the pictures differ widely in the extent to which they evoke food related stories. By far the most successful picture in this respect was F, which was constructed especially for this test. It was unique in two respects. It was the only picture that had food in it—i.e., a piece of meat—and it was the only picture which was

[5] The results on pictures G and I show identical trends but the frequencies are so small it was decided not to include them in the table.

'unstructured'—i.e., which was not a single scene, but three isolated picture elements which had to be combined into a plot. There was no other picture which approached it in any category except perhaps picture B in the D category.

In the second place, the data do not reveal any very great differences among the pictures in their sensitivity to changes in the hunger drive. For example, the increase in the D category from the one to the 16 hour group is about 13 percent for picture F and about 14 percent for picture B. The decrease in the G category for picture F is about 16 percent and for picture A about nine percent. The best pictures should be those which give enough food-related responses to be scored and which give the largest differential for different degrees of hunger. The size of the differential does not seem to vary markedly with the food-evoking capacity of the picture or with the nature of the picture, though obviously some pictures fail to evoke a certain characteristic at all. The rule seems to be that if a picture evokes a characteristic moderately frequently (in at least 15–20 percent of the *S*s), it will discriminate the two groups regardless of its other properties. It should be noted, however, that though this holds true for the size of the differential it does not hold true for its reliability. Only the differences for picture F approach statistical significance.

Furthermore there is an interesting exception to the rule. Picture E and picture A are both suggestive of satiation or at the very least of meal times (supper and lunch). The 16 hour group reacted to both pictures by producing less overall food imagery than the one hour deprivation group. If the data on F I for pictuers A and E are combined so that the percentage of *S*s who produced F I on either picture is recorded, the difference approaches significance (chi-square = 3.50). This suggests that with pictures implying goal activity there is a blockage of the normal food associations for hungry *S*s, a fact which is clearly in line with the decrease in their production of goal responses (G). The difference probably did not appear in picture F because the piece of meat was supposed to be fitted into the plot which almost required that food imagery be used by all *S*s.

The final fact to be gathered from Table V is that all the pictures except D showed a shift in mean NF score in the correct direction (more positive for the hungrier group). The discriminative power of the various pictures as reflected by the critical ratios of the differences was as follows: first, picture F (C. R. = 4.50); second, picture B (C. R. = 1.81); third, picture E (C. R. = 1.63); and fourth, picture A (C. R. = 1.00). Picture D showed small and insignificant, if consistent reversals in most categories and in NF score. It is difficult to draw any general conclusion from this except that, since it was the least food related of the five pictures, a certain minimal relation be-

tween a picture and the need it is supposed to reflect is absolutely necessary before it will accurately reflect that need.

Discussion

The apperceptive expression of a need.—The present experiment makes some important contributions to the central problem of this series of papers—namely, how a need expresses itself in perception and apperception. It has collected the first real data showing how thematic apperception stories are related to a known condition of the subjects. Murray (**9**) in his list of the ways in which story interpretations can be validated does not mention the technique used here. He is chiefly concerned with how an interpretation can be checked against biographical data, against the results of different tests, or against future behavior. Here the approach has been to discover *how* a need expresses itself so that the judge will have some actual basis for making or improving his interpretations of thematic stories.

Of course, judges have always used, consciously, or unconsciously certain clues for deciding whether or not a story or a series of stories indicated the presence of a need. The present results suggest that some of these clues may have been misleading. For instance, it has commonly been assumed (cf. **3**, p. 189) that phantasies serve the function of partially gratifying unfulfilled desires. One might therefore expect that hungry subjects would project more food objects which could serve to gratify an unfulfilled wish for food. The present data clearly indicate that this assumption (which Murray also makes—**10**, p. 260) is wrong, at least so far as these data are concerned. Or again, on the basis of Bellak's finding (**1**) that criticism increases aggressive words in such stories, one might suppose that hunger would increase the overall number of food related words and images, but it does not.[6] There must be many other such clues that clinicians use in making their interpretations, but not many of them have been made explicit. Since the only check on whether the clues used are correct is the checking of the total impression against some other source of knowledge, it is extremely difficult to improve the basis of such judgments because it is almost impossible to know what factors formed the total impression and to vary them systematically as they are validated against an outside criterion.

The present method promises better results. It can now be stated with some certainty that a person trying to diagnose hunger from thematic stories should look for: (1) food deprivation, both as the central plot of the story and as a secondary characteristic; (2) instrumental activities which successfully overcome the source of food deprivation but which often do not get the person food; (3) statements

[6] There is the possibility, of course, that the method of scoring used here—presence of food imagery in a story—is less sensitive than the method used by Bellak of counting every need-related word.

implying hunger or a need for food in the characters; (4) a fewer than normal number of references to goal activity such as eating, to friendly press such as an invitation to eat; and (5) a decreased responsiveness in terms of food imagery to pictures strongly implying satiation or the recent termination of eating.

These are the primary characteristics of 'hungry' stories. There are some secondary ones which are not as well established. For instance, the judge should not expect an increase in successful food getting activity, but he might look for an increase in anxiety over survival, and a decrease in thinking about food at an unreal level (wish phantasy or fulfillment).

It has been shown how these clues can be combined into a composite numerical score which gives a pretty good indication for most of the *S*s of how long they have been without food. Naturally it is hoped that this procedure will have a much wider significance than could be attached to the determination of hunger from *S*s' stories. It would be much simpler to ask the *S*s when they last ate! The results become of major importance only if changes occur in the same way in the same categories of response with other experimentally induced needs—particularly psychogenic ones. Other experiments in the series tackle this problem.

Theoretical considerations.—The data also shed some light on the nature of the processes involved in thematic apperception. The first point of theoretical interest is that the results are much the same as those obtained in the first experiment (7) on the effect of hunger on perception. Perception and apperception apparently respond to a need in much the same way. The chief difference lies in the greater freedom of the apperceptive processes to show a variety of changes under the conditions of the experiment. For instance, the first experiment showed an increase with hunger in the perception of objects instrumental to eating. The second experiment elaborates this finding by showing that instrumental activities as well as objects are concerned but the activities are not so much connected with food getting as with eliminating deprivation. One might even argue that the *S*s in the first experiment saw plates rather than oranges, not because the plates were instrumental, but because they were empty. That is, the increase in perception of food-related objects rather than actual food may have signified not instrumentality but the deprivation so characteristic of the 'hungry' stories.

Or again, in the first experiment, the *S*s failed to perceive an increased number of food objects, but in the second experiment, the more sensitive apperceptive processes reflected an actual significant *decrease* in goal activity. Also, perception did not clearly differentiate the four hour from the 16 hour deprivation group. Apperception did, though it did not reflect many significant changes between the one and four hour groups.[7]

[7] The curious increase in the estimated numbers of neutral objects in the first experiment has not been checked here by analyzing the stories for characteristics unrelated to food. It is planned to do this in the future.

Even if it is granted that apperception is merely reacting more sensitively than perception but according to the same underlying process, the problem remains as to what the process is and why it functions in this particular way. Cattell suggests (3, p. 180) that there are three processes involved in what is ordinarily called 'projection.' The first, 'naive inference from limited personal experience,' could be illustrated in the present experiment by the *Ss*' 'naively' inferring that people in the pictures have been deprived of food because they have been. The second, 'unconscious, immediate, or true projection,' [8] can probably not occur as a defense mechanism against such a non-ego-involving drive as hunger, but a case could be made for the fact that the hungty *Ss*' tendency to attribute the hunger motive to others in the pictures objectifies the need so that it can be handled more effectively. The third, 'projection of press required by emotional state,' [8] is illustrated perhaps by the refusal of the *Ss* to give goal activity responses to situations normally evoking them. That is, since they feel hungry, they see the world as a place with less food in it which 'explains' how they feel.

The data indicate that a number of such processes occur and the writers feel that there is little to be gained by trying to label certain processes projection and others something else. It is all behavior and should be treated like any other behavioral attempt at adjustment by the organism. That is, perception and apperception are determined like any other behavior is—partly by the state of the *S* (including needs and drives), partly by his past experience, and partly by the objective situation. The major difference between this and any other situation is that the objective cues have been reduced or nearly eliminated so that the behavior must be determined mostly by needs and past experience.[9]

The view that the cognitive and perceptual processes involved here do not function essentially differently from any behavioral attempts at adjustment is supported by Knapp's (6) and Allport and Postman's (1) similar conclusions from studies of rumor and by the nature of the changes that actually occurred here. That is, in general it is to the obvious advantage of the organism to concern itself more and more with instrumental rather than goal activity as the need increases. It is also useful to emphasize the problem (deprivation) and ways of overcoming it (I d+). Perhaps the only surprising element in the situation is that this mobilization at the imaginative level of the resources of the organism toward solving or handling the problem of food deprivation goes on apparently very largely without the *S*'s knowledge or conscious intent.

Secondary implications.—The data bear on several problems that have been raised insistently by previous research in this field. For instance, what is the effect of recent experiences on the validity of projective tests such as

[8] The distinction between this and the third process is essentially the same as Murray makes (8) between supplementary and complementary apperceptive projection.

[9] The difference between this and the situations in which Freud described projection is that he was dealing with cases in which the need was so strong as to render ordinary objective determinants relatively unimportant. In phantasy situations the objective determinants are rendered unimportant by the experimenter so that behavior can be more completely reflective of needs. The end result is often the same except that the Freudian projection will only occur in cases of such extreme need that the ordinary objective determinants of an interpretation are overridden.

the TAT? As Coleman says, "the TAT should not be subject to such day-to-day fluctuations" (5, p. 257). Both he and Bellak (2) conclude that it is not—at least not at any very fundamental level. The present data do not support this conclusion entirely. If, for instance, an *S* takes a TAT at 11 A.M. in the morning and he has missed his breakfast (a situation comparable to the 16 hour deprivation group), there is much greater likelihood of his telling a general deprivation story to picture A (the 'beachcomber' from an early TAT edition) than if he takes the TAT shortly after lunch. This does not mean, of course, that the basic picture of the man's personality structure will be altered. There may be alteration only in the manifest and not in the latent content, as in Freud's dream analyses. But it does suggest caution at least to the extent of finding out what significant things have happened to a person in the last few hours or even days before administration of a projective test.

What kinds of pictures are most effective in eliciting the needs of the *S*? Here again there have been several studies (12, 13) most of which conclude that they should be vague and contain characters with whom the person can identify. The basis for choosing these characteristics has not been any demonstrated discriminative power of such pictures but rather the quality of the stories they produce. So far as any conclusion can be drawn from the few pictures used here, it would seem that the projection pictures should contain objects or situations related to the need one is interested in measuring. This is to insure a sufficient frequency of responses related to the need under consideration to get differential reflections of it. To be more explicit, the pictures should contain goal objects in an unstructured relationship to a person with whom the *S* can identify. In this connection, the technique of presenting three or more elements which the *S* must weave into a story seems particularly promising. They should also present situations suggestive of terminal or goal activity and of deprivation or blockage of instrumental activity. The present set of TAT pictures is particularly deficient in the first characteristic. That is, they do not contain goal objects around which themes can be woven, but present instead situations which may be so ambiguous that like picture D in the present series they fail to evoke much of any response (and hence no differential response) related to the needs which the experimenter is interested in measuring. These points, it must be emphasized, are merely suggestions which arise from a very limited number of pictures under the influence of a simple physiological drive, and yield no firm factual support for wide generalization. However, they have seemed worth making, since this is one of the first attempts which has been made to determine the effectiveness of pictures in terms of their actual discriminating power for the need being measured.

Summary

A group of 81 men, applicants for admission to Naval submarine training school, wrote out brief stories in response to seven projected pictures, 38 after one hour, 21 after four hours, and 22 after 16 hours of food deprivation. Results from 22 more men who wrote stories in

response to five of the seven pictures were used in the analysis of the comparative merit of the pictures in eliciting responses reflecting hunger. The stories were analyzed by two judges independently for 23 characteristics relating to the hunger drive. The following conclusions appear justified:

1. The judges were able to agree in categorizing story content in a large percentage of cases. The correlation between the categories for two groups of 19 *Ss* under the same condition was .89, indicating a considerable degree of stability for group percentages in various categories.

2. As hunger increased, there was no overall increase in the percentage of *Ss* showing food imagery or food themas, but there was a decided increase in the percentage showing food deprivation themas, characters expressing a need for food, and activity successful in overcoming deprivation, but not always instrumental in getting food. On the other hand, as hunger increased, there was a decided decrease in the amount of goal activity (eating) and in friendly press, favorable to eating.

3. A composite need food score was devised by scoring +2, +1 or −1 all instances in each record of categories which showed reliable increases or decreases for the groups as hunger increased. The mean need food score differentiated reliably the three deprivation groups with little overlap in score between the one and 16 hour deprivation groups.

4. No outstanding differences in sensitivity to reflecting hunger was discovered for the various pictures used provided the picture was enough related to hunger to produce a sufficient number of food-related responses to show a differential. The most successful picture in differentiating the three food deprivation groups was one containing a goal object (food) in a completely unstructured relationship to a young man with whom the *Ss* could identify.

The application of these findings to the clues used in interpretation of Thematic Apperception Test records is discussed. It is suggested in particular that the amount of need deprivation and of instrumental activity present in stories is a better index of the strength of a need than is the amount of goal activity. If the results are confirmed at more complex need levels, it should be possible to obtain a single score measuring need strength which would combine the shifts occurring in story content as the strength of the need increases.

Consideration of the theoretical implications of these results indicates that it is desirable to treat apperceptive behavior as functioning like any other type of behavior rather than to attempt to fit it

under the limited principles governing Freudian or defensive projection which appears to be a special case.

(Manuscript received November 18, 1947)

1. ALLPORT, G., & POSTMAN, L. *The psychology of rumor.* New York: Holt, 1947.
2. BELLAK, L. The concept of projection. *Psychiatry*, 1944, **7**, 353–370.
3. CATTELL, R. B. Projection and the design of projective test of personality. *Char. & Pers.*, 1944, **12**, 177–194.
4. CLARK, R. M. A method of administering and evaluating the Thematic Apperception Test in group situations. *Genet. Psych. Monogr.*, 1944, **30**, 3–55.
5. COLEMAN, W. The Thematic Apperception Test. I. Effect of recent experience. II. Some quantitative observations. *J. clin. Psych.*, 1947, **3**, 257–264.
6. KNAPP, R. H. Experiments in serial reproduction and related aspects of the psychology of rumor. Unpublished Ph. D. thesis, Harvard University, 1948.
7. MCCLELLAND, D. C., & ATKINSON, J. A. The projective expression of needs. I. The effect of different intensities of the hunger drive on perception. *J. Psychol.*, 1948, **25**, 205–222.
8. MURRAY, H. A. The effect of fear upon estimates of the maliciousness of other personalities. *J. soc. Psychol.*, 1933, **4**, 310–329.
9. MURRAY, H. A. Techniques for a systematic investigation of fantasy. *J. P ychol.*, 1937, **3**, 115–143.
10. MURRAY, H. A. *Explorations in personality.* New York: Oxford Univ. Press, 1938.
11. RODNICK, E. H., & KLEBANOFF, S. G. Projective reactions to induced frustration as a measure of social adjustment. *Psychol. Bull.*, 1942, **39**, 489.
12. SARGENT, H. Projective methods, their origins, theory and application to personality research. *Psychol. Bull.*, 1945, **42**, 257–293.
13. SYMONDS, P. M. Criteria for the selection of pictures for the investigation of adolescent fantasies. *J. abnorm. soc. Psych.*, 1939, **34**, 271–274.

STUDIES IN SPATIAL LEARNING. VI. PLACE ORIENTATION AND DIRECTION ORIENTATION

BY BENBOW F. RITCHIE[1]

Swarthmore College

Introduction

In the first and second articles in this series (**3, 4**), it was pointed out that the purpose of this series was to develop some of the important implications of Tolman's 'expectancy theory.' It was also pointed out that when this theory is applied to problems of maze learning it is very similar to Woodworth's 'place-learning hypothesis' (**6, 7**). Both of these theories or hypotheses emphasize the idea that what is learned in any spatial problem is not simply a specific response sequence, but is *something* about the spatial relations between the various objects in the rat's environmental field. Neither of these hypotheses, however, specify *what* is learned about these spatial relations, and because of this lack of specificity these hypotheses are not sufficiently precise to permit experimental investigation.

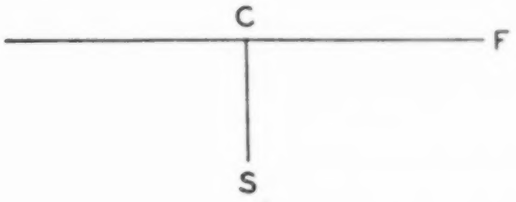


Fig. 1. Schematic drawing of simple T-maze

Consider now a rat which has been trained on a simple T-maze (see Fig. 1) so that whenever the rat is hungry and is put at S he runs quickly up to C and turns without hesitation to the right towards F. Here is an extremely simple example of spatial learning. Now how do Tolman's and Woodworth's theories of such learning differ from what has been called 'the orthodox S-R theory of maze-learning'? Primarily the difference between their theories and the S-R theory seems to be that the latter kind of theory regards such learning as the acquisition of a single disposition, namely, the disposition to take the

[1] This investigation is part of a co-ordinated research program of the Psychological Laboratory at the University of California. The writer is deeply indebted to Professor E. C. Tolman for generous advice and aid. The present paper is a portion of a dissertation presented to the Faculty of the Graduate School of the University of California in partial fulfillment of the requirement for the degree of Doctor of Philosophy.

practiced path, SCF. Tolman and Woodworth, on the other hand, suggest that not only is this disposition acquired as a result of the training, but other dispositions as well. Now, whether other dispositions are acquired during training on a particular path is, of course, an empirical question, and it is the primary purpose of this series of studies to answer this question, and to discover something about the nature of such *auxiliary dispositions* if they do in fact exist.[2]

In the present study we will confine our discussion to two kinds of auxiliary dispositions, (1) an auxiliary place disposition, and (2) an auxiliary direction disposition. We will first consider the meaning of these terms, and then discuss the evidence that we have for believing that such auxiliary dispositions are sometimes acquired.

1. *Auxiliary place dispositions.*—If after training on a simple T-maze, (1) the path from S to C is blocked (see Fig. 2) and (2) several

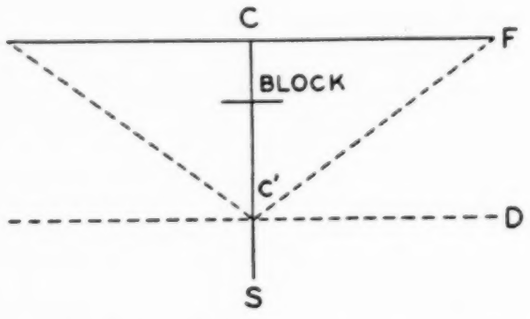


FIG. 2. Schematic drawing of test situation for auxiliary dispositions. S = original starting place. C = original choice point. F = original food place. C′ = new choice point.

new paths are offered to the rat, one of which (i.e., C′F) points directly to the original food place, *then*, if under these changed conditions the rat chooses path C′F, we shall say that the rat has acquired an auxiliary place disposition.

2. *Auxiliary direction dispositions.*—If under exactly the same conditions, the rat then chooses a path (i.e., path C′D) which runs parallel to the original true path (i.e. path CF) and points in the same direction, we shall say that the rat has acquired an auxiliary direction disposition.

The experiment reported in the first article in this series (3) presented evidence that auxiliary place dispositions are sometimes acquired by rats.

[2] I shall call the disposition to take the practiced path 'the primary disposition,' and all others 'auxiliary dispositions.' It should be emphasized however that these terms are merely introduced for the sake of abbreviation, and not because I believe that the disposition to take the practiced path is more fundamental than the other dispositions. In fact it may be discovered that the auxiliary dispositions are more basic than the primary disposition in the sense that the acquisition of the auxiliary ones is a necessary condition for the acquisition of the primary disposition.

Briefly the procedure was this. A group of 56 maze-wise rats was trained in a new spatial problem which involved running over a table-top, through an alley, and then down an elevated path which ran at right angles to the starting path. When these rats had received five such training trials, the original path was blocked at the end of the alley, and 18 elevated paths were attached to the circular table-top radiating from its edge at intervals of 10 degrees. On the test trial *all* of the rats ran down the alley to the block, and then returned to the table top. They all explored the table-top and the initial segments of the attached paths. After a great amount of such exploration all but three of them chose one of these new paths and ran out to the end of the path. Thirty-six percent of the rats in this group chose the path which pointed most directly towards the *former* food location. The rest of the rats in the group were distributed randomly on the remaining 17 paths. We concluded from these results that 36 percent of our rats had acquired an auxiliary place disposition, and we called this disposition 'an expectation of food at location L.'

The evidence for the existence of auxiliary direction dispositions is not as simple or direct as the evidence for auxiliary place dispositions. In fact, it was only after observing the behavior of some rats in the first experiment reported in the third article in this series (5) that we hit upon the idea of auxiliary direction dispositions. In this experiment we observed that the rats in the control group after having been trained on a simple T-maze, continued to run in the same direction (i.e., towards the west wall) no matter where the maze was put in the room, as long as the path from the choice point to the food (i.e., C′F) was parallel to the original CF path. As a consequence we concluded that this behavior *might* be the result of having acquired an auxiliary direction disposition.

This, then, is all the evidence we have at present for the existence of such direction dispositions. Although this evidence is far from overwhelming, it was sufficiently cogent to suggest that further research on this point might be fruitful.

We turn now to a new problem. It should be clear that if such auxiliary place or direction dispositions are sometimes acquired by rats, in such cases the behavior of the rats is a function not only of the intra-maze environment, but of the extra-maze environment as well. Furthermore the extra-maze cues in such cases seem to play a peculiar role, a role which is different from the role usually attributed to the more proximal intra-maze cues. Thus, consider a rat on the test trial in which he selects a completely new path (one which points towards the original food place, or one which points in the same direction as the original true path). In such a case the extra-maze cues are 'received proximally' from a wholly new perspective, and yet somehow they seem to be recognized and interpreted in a way that is appropriate to the objective spatial relation of the objects in the environmental field. By this I mean that the rats are able to choose new paths which point towards the same place in the objective environment, or point in

the same direction, despite the fact that the perspectival relations of the rats to the extra-maze cues are entirely novel.

The question now arises as to whether this 'grasping of objective spatial relations' (despite changes in the immediate proximal perspective) is restricted to the original, practiced starting place or starting path, or whether when acquired in one starting place the auxiliary direction or place dispositions are of a sort to generalize to new starting places in the objective environmental field. The following experiment was designed to answer this question. First, a group of rats was trained on a simple T-maze from a single starting place, and then after this maze had been learned, we determined what auxiliary place or direction dispositions the rats had acquired, if any, from a new starting place. For reasons which will be explained in the conclusion of this article, the *new* starting place which was used was a place on the opposite side of the choice point of the T-maze, on the same axis as the original starting path, and at the same distance from the choice point as the original starting place.

Animals

Fifty pigmented male rats approximately three months old were used in this experiment. These rats were MxM's drawn from the Tryon stock in the rat colony of the University of California. None of them had had any previous maze training. They were put on a 24-hour wet-food maintenance schedule five days before the beginning of the preliminary training. During the training they were run between 8 and 11 P.M.

Apparatus

Figs. 3 and 4 present diagrams of the apparatus which were used. In Fig. 3 we see the apparatus used in the preliminary training. It consisted of an unpainted wooden circular table

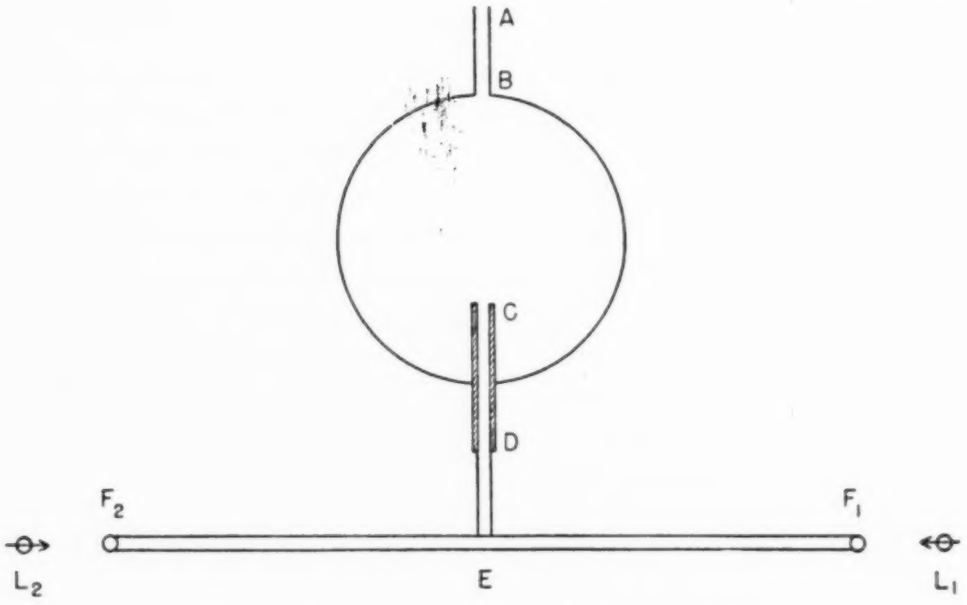


Fig. 3. Apparatus used in preliminary training

top 91 cm. in diameter and several elevated wooden paths painted black. The path AB was 27 cm. in length and was used as a starting path. The path CE was 75 cm. in length, while path F_1F_2 was 2.5 meters long. At each end of the F_1F_2 path were small paper food cups three cm. in diameter and three cm. in depth. Twenty cm. beyond each end of the F_1F_2 path were small lights L_1 and L_2, held in a socket so that the beam of light was directed on the food cup below. These lights contained 15-watt bulbs which were covered with the paper carton in which they were sold. They were 20 cm. above the level of the F_1F_2 path. The only other illumination in the room was located at L_3 (see Fig. 5A). At this place was an ordinary desk lamp containing a 75 watt bulb. The reflector of this lamp was turned in such a way that its light was directed on the white wall of the room at an angle of 45 degrees. Thus, this corner of the room (see Fig. 5A) was brightly lighted while the rest of the room was in deep shadow. Fastened to the sides of the path DE were two pieces of unpainted plywood, which were 47 cm. in height and 47 cm. in length. These formed an alley which began 20 cm. from the center of the table-top and ended 27 cm. away from the choice point E.

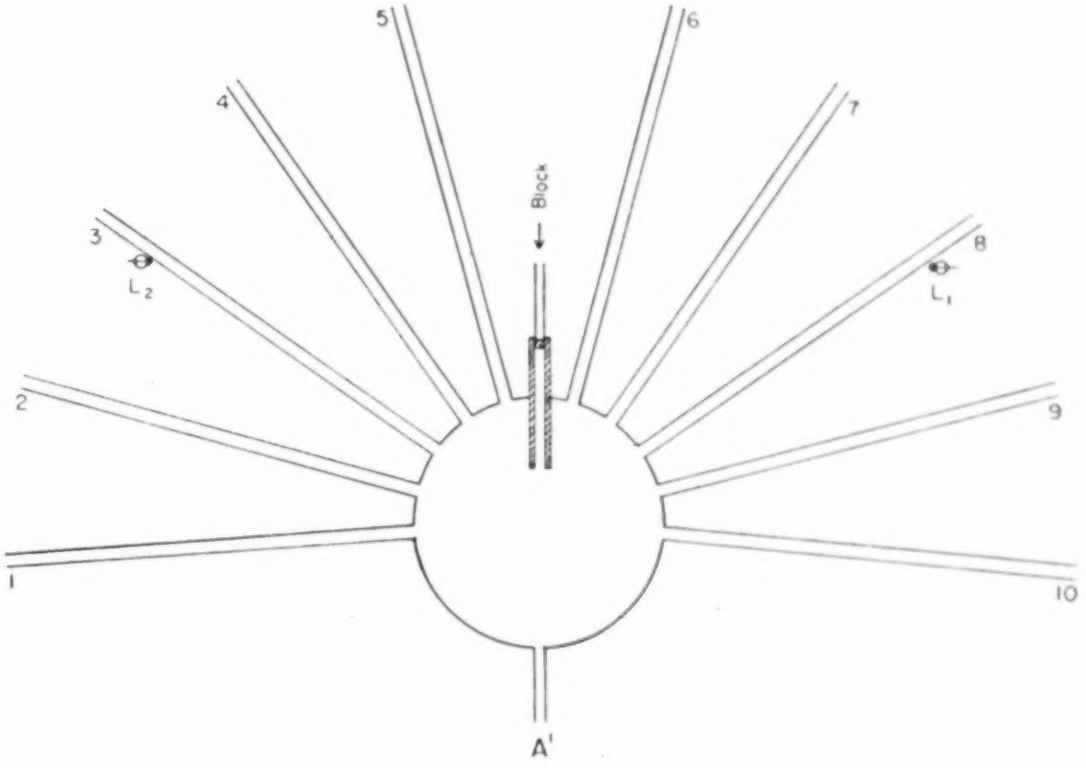


FIG. 4. Apparatus used in test trial

In Fig. 4 we see the apparatus used in the test trial. This consisted of the same starting path, circular table-top, alley on path CE, and lights at L_1, L_2, and L_3. But the path F_1F_2 with food cups was removed, and a block was placed at the end of the alley on path CD. Then 10 paths 1.5 meters in length were placed around the table-top. These paths began at a point 95 degrees to the right of path CE and radiated in a counterclockwise fashion, each path being placed 20 degrees to the left of its neighbor.[3]

As can be seen in Fig. 5A the choice point E of the maze used in the preliminary training was in the center of the room and the F_1F_2 path ran parallel to the north and south walls. The apparatus used in the test trial was set exactly opposite to the other apparatus (see Fig 5B) so that the starting point A′ was now the same distance from the south wall as it had been from the north wall in the preliminary training. The positions of the lights remained however unchanged.

The rats lived in cages in another room and were carried in their cages from their living room to the experimental room for their training. Their cages were then put in a rack on the same axis

[3] There is one exception to this. Path 5 was 30 degrees to the left of path 6.

as that of the starting path and the choice point of the maze used in the preliminary training (see Fig. 5A). On the test trial they were put in the same rack. The *E* carried the rats by hand, put them on the apparatus by hand, and then sat on a stool beside the rack where the rats were kept.

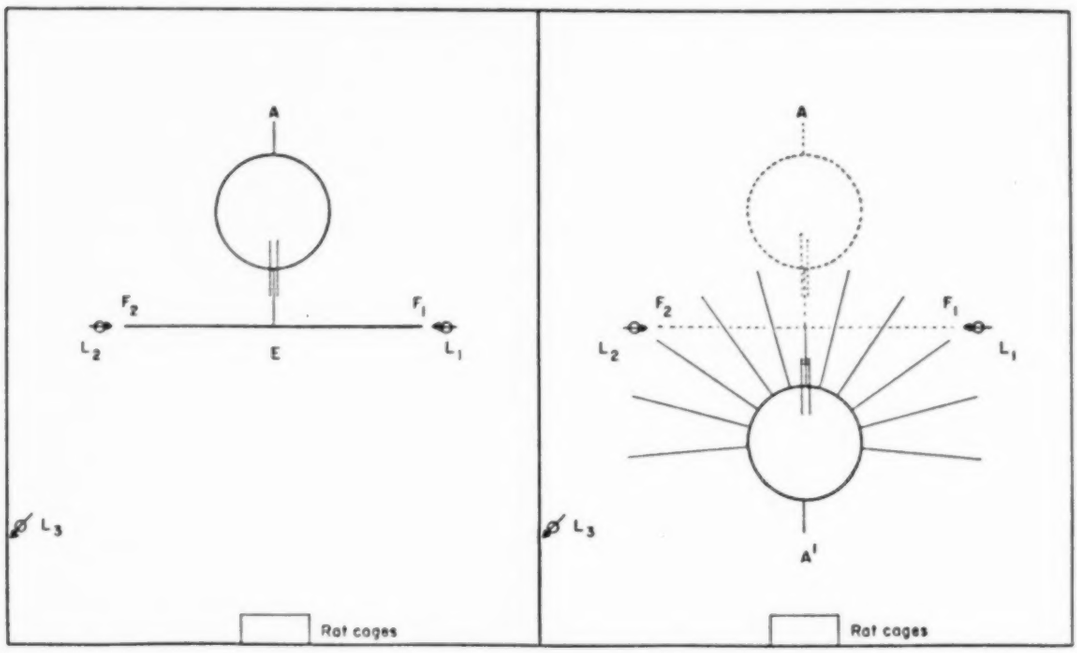


FIG. 5. Floor plans showing positions of apparatus. Positions during preliminary training in 5A. Positions during test trial in 5B.

METHOD

Pre-test procedures.—On Days 1 and 2 the rats were trained in another room to accustom them to turning on straight paths and to eating from paper food cups.

On Day 3 a block was placed at D on the apparatus in Fig. 3 and the rats were given six trial of running from F_1F_2.[4] On each of these trials the rats were allowed to eat at F_1 for 20 sec. and then were returned to the cage in the rack.

On Day 4 the rats were given two trials from D to F_1 (with a block at D) and two more trials from D to F_1 (without the block at D). Then they were give two trials through the alley from C to F_1.

On Day 5 they were given two more trials through the alley from C to F_1. They were then given four trials from the start A to F_1.

On Days 6 and 7 they were given eight daily trials from A to F_1. Thus by the end of training on Day 7, each rat had run 20 times from A to F_1.

Test procedures.—On Day 8 one test trial was given. The apparatus was changed to that represented in Fig. 4, and placed in the room as pictured in Fig. 5B. Each rat was started at A' and allowed to explore the blocked alley, the table-top, and the various alternative paths which radiated from it. If any rat put both front feet on any path but did not run out more than two feet on the path[5] this was recorded as a 'partial choice.' A rat's trial ended as soon as it had chosen one of the paths and had run out on it at least four feet. If any rat took more than seven min. to make such a choice, then it was removed.

---

[4] Twenty-three of the rats were trained to find food at F_1. This group will be the 'F_1 group.' Twenty-seven of the rats were trained to find food at F_2, and they were called the 'F_2 group.' In the following description of the training procedure we will consider only the F_1 group. The analogous kind of training was of course given to the F_2 group.

[5] It was discovered in the first short-cut experiment (3) that any rat which ran more than two feet out on a path ran all the way out. The same behavior was observed in this experiment.

Results

On the test trial, eight of the 50 rats were discarded because they failed to make a choice after seven min. After having explored the paths and table-top, all of these eight rats returned to the center of the table and refused to move from there except to return either to the alley or to the starting path.

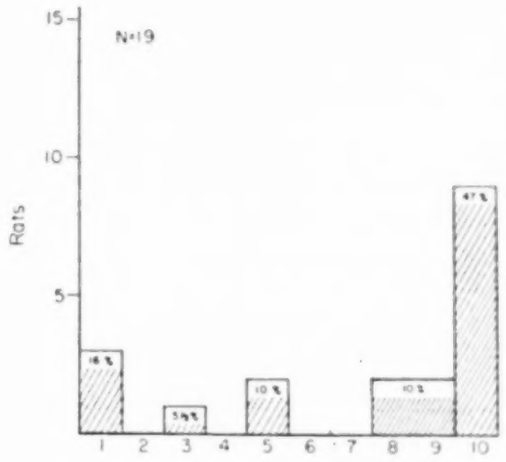


Fig. 6. Percentage of path choice in the F_1 group

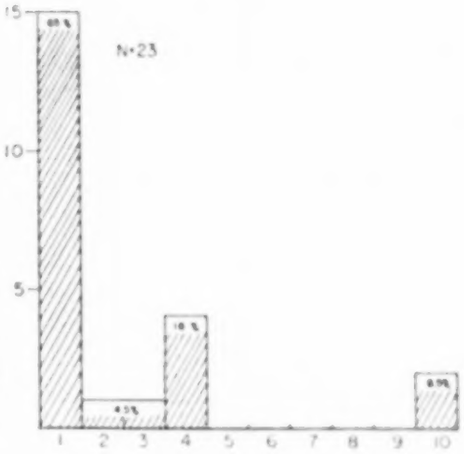


Fig. 7. Percentage of path choice in the F_2 group

Fig. 6 shows the distribution of choices made by the remaining rats in the F_1 group. Inspection of this distribution shows that only 10.5 percent of the rats in the F_1 group chose path 8, the one which points most directly towards F_1. The distribution of the choices in the F_2 group (see Fig. 7) is similar; only 4.5 percent of the rats in this group chose path number 3, the one which points towards F_2.

While only a few of the rats chose the most direct path to the original food place, the path chosen most frequently by both groups was the last path on the side of the table which faced the original food place. Thus, 47 percent of the F_1 group chose path number 10, and 65 percent of the F_2 group chose path number 1.

Now, since the problem set for the F_1 group was the mirror image of that set for the F_2 group, the choices of both groups can be combined as in Fig. 8, where A is the choice of the last path on the side facing the former food place (e.g. in the F_1 group path 10, and in the F_2 group path 1), C is the most direct path to the former food place, and so forth. When the results of both groups are combined in this way we see that 57 percent of the 42 rats chose the last path on the food side, while the distribution on the other paths never exceeded 11 percent.

In Fig. 9 we see the distribution of partial choices (i.e. in which the rats did not run out more than two feet) of the F_1 group, and in Fig. 10

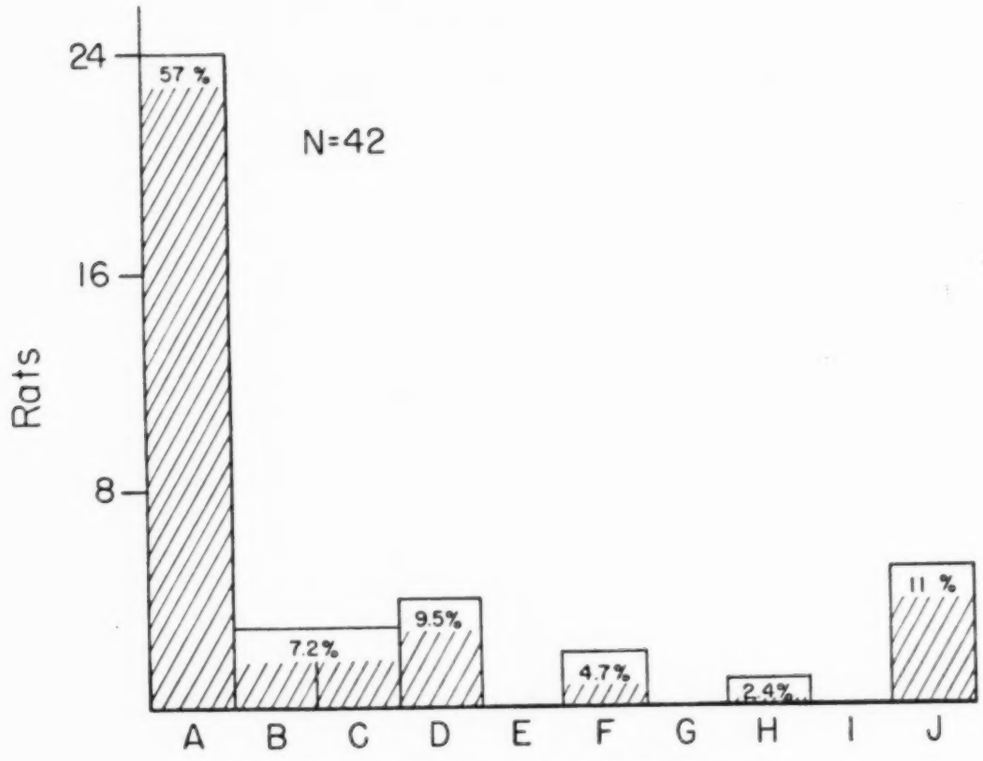


FIG. 8. Percentage of path choice in both groups

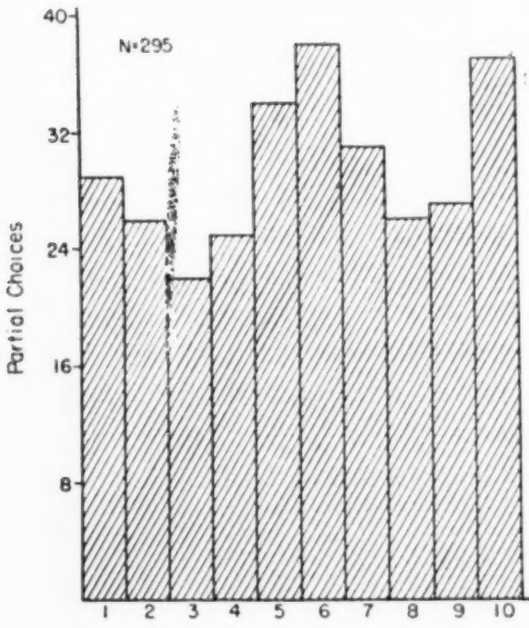


FIG. 9. Partial choices in F_1 group

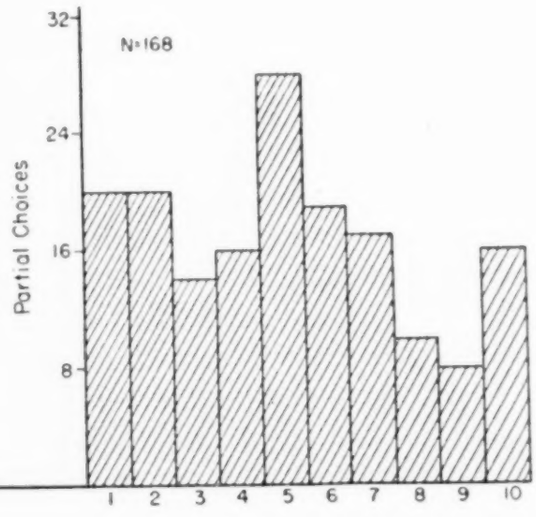


FIG. 10. Partial choices in F_2 group

we see the distribution of partial choices for the F_2 group. These two distributions show that the rats in both groups explored all paths about equally before making a choice. There was, however, one difference between the two groups. The F_1 group made approximately twice as many partial choices as the F_2 group. The mean number of partial choices for the F_1 group was 15.5 with a standard deviation of

8.5, while the mean number of partial choices for the F_2 group was 7.2 with a standard deviation of 5.5. The critical ratio of the difference between these two means was 11.2.

Conclusion

The results of this experiment show that auxiliary direction dispositions sometimes generalize to new starting places in the objective environmental field. Further, it is clear that in *our* experimental situation auxiliary place dispositions do not. We cannot, however, conclude from these results that auxiliary place dispositions never generalize to new starting places, since we know nothing, as yet, about the environmental factors which favor place as compared with those factors which favor direction dispositions.

Far from being unexpected, results such as these were predicted from our interpretation of the results of other experiments in this series. In fact, the notion that some kind of auxiliary disposition was generalized to other starting places had been suggested by the results of the early Honzik and Tolman (2) experiment, 'The perception of spatial relations by the rat.' In this experiment rats were trained on an elevated T-maze (Fig. 11) to prefer the shorter path from S_1 to F_1 and to avoid the longer path S_1 to F_2, although there was food at both F_1 and F_2.

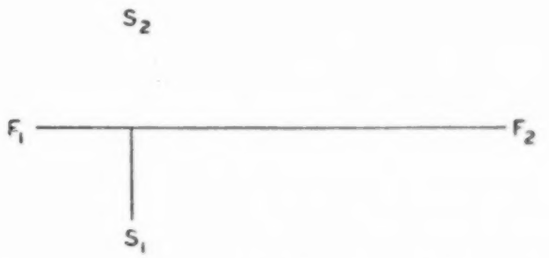


Fig. 11. Schematic drawing of maze used by Honzik and Tolman (2) in which rats were first trained to prefer the shorter path S_1 to F_1. Then, after having been shocked at F_1, they were tested from S_1 and from S_2.

After this preference was established, the rats were carried by hand and shocked in the food box at F_1. When they had recovered from this shock they were again started at S_1 and 87 percent of the animals which had been going consistently to F_1, now turned and went to F_2. Next, in order to demonstrate that the shock had not merely served to reverse their position or turn habit, another group of animals were given exactly the same training, and then after being shocked at F_1, they were tested from a *new* starting place S_2 (see Fig. 11). Again 80 percent of the rats went to F_2. Such results strongly suggest that some kind of auxiliary disposition which is acquired when the rat is practiced from S_1, generalizes or transfers to the new starting place S_2.

A simpler example of apparently the same phenomenon is provided by the experiment in the second article in this series (4). Here two groups of

rats were trained on the simple elevated maze shown in Fig. 12. Both groups of rats were hungry and were given six trials a day, three starting from S_1 and three starting from S_2. One group (the Place Group) were required to learn to go always to the same place, for example to F_1, while the other group (the Response Group) were required to learn to turn always

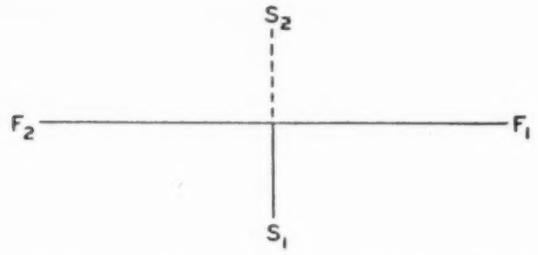


FIG. 12. Schematic drawing of maze used by Tolman, Ritchie, and Kalish (4)

to the right at the choice point. Thus, when started at S_1 members of the Response Group were required to go to F_1 and when started at S_2 to go to F_2. If auxiliary place or direction dispositions (1) transfer to new starting places, and (2) support or reinforce the related primary dispositions or dispositions to take the practiced path, then the problem for the Place Group should be easier than the problem for the Response Group.[6] The results of this experiment showed that place-learning in a situation where there were marked extra-maze cues was much easier than response-learning.

But such results do not provide unambiguous evidence for the transference of auxiliary dispositions, since other hypotheses can account for the results equally well. For example, it might be argued that in such a situation it is difficult to learn to make the same turn response repeatedly, since a situation like this makes for response variability. Thus, the Place Group which is permitted variability learns more repidly than the Response Group which is not.

In order to eliminate the possibility of such an explanation, the experiments in the third article in this series (5) were performed. In the second experiment in this third article, the Experimental Group was given the same kind of training as the Place Group in the experiment just described. The Control Group, on the other hand, was given two separate place problems. They were trained to run from S_a in one room (Room A) to F_a, and to run from S_b, in another room (Room B), where the extra-maze cues were totally different, to F_b. When started in Room A they were required to turn right, and in Room B they were required to turn left. Thus the pattern of turn responses in the experimental and control groups was identical, and the only difference between the problems set for the two groups was that the experimental group were required to go to the same food place from two different starting places, while the control group had to go to two

[6] This conclusion is of course based on the further assumption that in situations where there are marked extra-maze cues, turn dispositions (e.g. to the right) do not generalize to new starting positions. As yet we have no clear evidence that such turn dispositions are ever acquired in such situations. Blogett and McCutchan (1) suggest that when extra-maze cues are eliminated such turn dispositions may be acquired.

different food places from two different starting places. If our interpretation of the previous experiments has been correct then the one-place or experimental group should learn faster than the two-place control group. This hypothesis was confirmed by the results of this experiment. Finally this article concluded with a theoretical explanation of these results in terms of the transference of a short-cut disposition (i.e. the auxiliary place disposition).

The experiment in the present article was performed to test this theoretical explanation, and the results indicate that although there is no evidence that auxiliary place dispositions transfer to new starting places, there is good evidence that auxiliary direction dispositions do.

Summary

1. Orthodox S-R theory has seemed to argue that the only habit that is learned in a maze problem is a disposition to take the practiced path. Tolman's and Woodworth's theories, on the other hand, emphasize that other, auxiliary dispositions are also acquired.

2. Two kinds of auxiliary dispositions are herein distinguished, (1) place dispositions and (2) direction dispositions, and the evidence for the existence of such dispositions is reviewed.

3. To determine whether such auxiliary dispositions generalize to new starting places or paths, a group of 50 rats was trained on a simple T-maze similar to the one used in the short-cut experiment in the first paper in this series (3). After this training was completed, each of the rats were started from a place which was on the opposite side of the choice point of the original maze, and on the same axis as the original starting path.

4. Fifty-seven percent of the rats exhibited an auxiliary direction disposition. That is, they chose the last path on the side facing the former food place. There was no evidence that these rats also acquired an auxiliary place disposition. Only 7.2 percent of the animals chose the path that pointed directly towards the former food place.

(Manuscript received November 15, 1947)

References

1. Blodgett, H. C., & McCutchan, K. Place versus response learning in a simple T-maze. *J. exp. Psychol.*, 1947, 37, 412–422.
2. Honzik, C. H., & Tolman, E. C. The perception of spatial relations by the rat: A type of response not easily explained by conditioning. *J. comp. Psychol.*, 1936, 22, 287–318.
3. Tolman, E. C., Ritchie, B. F., & Kalish, D. Studies in spatial learning. I. Orientation and the short-cut. *J. exp. Psychol.*, 1946, 36, 13–24.
4. Tolman, E. C., Ritchie, B. F., & Kalish, D. Studies in spatial learning. II. Response learning versus place learning. *J. exp. Psychol.*, 1946, 36, 221–229.
5. Ritchie, B. F. Studies in spatial learning. III. Two paths to the same location and two paths to two different locations. *J. exp. Psychol.*, 1947, 37, 25–38.
6. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.
7. Woodworth, R. S. Reenforcement of perception. *Amer. J. Psychol.*, 1947, 60, 119–124.

AN ATTEMPT TO RECONCILE UNLEARNING AND REPRODUCTIVE INHIBITION EXPLANATIONS OF PROACTIVE INHIBITION[1]

BY B. R. BUGELSKI

University of Buffalo

In a previous experiment (1) in which *Ss* were required to learn four lists of paired-associate nonsense syllables in rapid succession, it was noted that a curious relation seemed to hold for the learning scores of the four different series. The first and third sets were learned relatively well, the second and fourth sets poorly. This relationship suggested a possible method for testing the Factor X (unlearning) hypothesis put forward by Melton (4) as an explanation of some proactive and retroactive inhibition effects. The unlearning hypothesis, as McGeoch (3) has tried to show, is not in essential conflict with extinction or reproductive inhibition explanations, but as originally presented, it was based largely on negative evidence, that is, on the absence of sufficiently numerous 'intrusion' errors in the recall period of typical retroactive inhibition experiments. In his interpretation of unlearning in terms of competition of responses, McGeoch indicated that such competition would have to occur during interpolated learning in retroactive inhibition experiments, but beyond his statement that "overt intrusions from original to interpolated lists are known to occur," very little appears to be known concerning the nature or frequency of such intrusions. One of the reasons for the paucity of data on intrusion errors is that experimenters usually record errors and successes by simply checking instead of obtaining a verbatim record of all responses made by *Ss*. Again, *Ss* are prone to sit silently and actively inhibit possible intrusive associations when such present themselves because of the self-instructed set to give only correct responses. There appears to be no good reason why *Ss* should not be encouraged to report the various associations that occur even if they recognize them to be incorrect. By such encouragement, more evidence concerning the nature of the competitive process in the learning situation should be obtained without sacrifice of major objectives. The experiment to be described made use of such deliberate exploitation of *Ss*' errors in the attempt to answer the question raised by

[1] The major portion of this paper was read at the Detroit meetings of the American Psychological Association in September, 1947.

McGeoch as to whether unlearning is, in fact, different from the competition he believed responsible for reproductive inhibition.

The present experiment was designed to observe the interference effects of learning and presumed unlearning of a series of materials in the framework of an extended proactive inhibition experiment. *Ss* were asked to study one list of paired-associate nonsense syllables after another for 10 successive lists. With equal time periods allowed for each list, it is argued that, according to the unlearning hypothesis, all other factors being equal, the second list should not be learned as well as the first because some of the time alloted must be consumed in unlearning List I. Because of this consequent poorer learning, List II will be unlearned more easily in the time alloted for List III, so that List III should be better learned. This process should continue indefinitely and all odd-numbered lists in any series should be better learned than the even-numbered series. The literature on learning in such situations is meager, although at least two studies report suggestive data. Underwood (6) had *Ss* learn up to six series of paired associates and found the predicted drop on the second series of paired associates and a rise on the third. His fourth series, however, showed a further increase which he was unable to explain except in terms of a growing discrimination factor such as Gibson (2) emphasizes. Underwood was perplexed by the improvement in List III inasmuch as he expected a general downward trend in view of the earlier results of Twining (5) who found a progressive decrease through five successive series. In the development of the present experiment, it was not expected that the results would show a systematically cycloid effect of decline and improvement in absolute scores because the practice effect and development of discrimination might result in a steady rise in learning scores; on the other hand, the cumulative frustration or proactive effect might result in a general decline. Such general trends as progressive rise and decline are probably due to the difficulty of the materials and sophistication of the *Ss*, and would not invalidate the present interpretation if it could be shown that the rate of change followed the predicted order. A smaller loss on odd-numbered series, or a smaller gain on even-numbered series should be sufficient to at least illustrate the unlearning hypothesis. The experiment as so far described, even if it demonstrated the superiority in learning of the odd-numbered lists, would only indirectly indicate that unlearning was occurring in the even-numbered lists, without, in any way, indicating the nature of such unlearning. By asking the *Ss* to respond to every stimulus correctly, if possible, and incorrectly with the first association (in more or less free association style), it was hoped that additional light would be cast on the nature of the unlearning process. In view of the instructions given, the nature and frequency of intrusive errors, their persistence if any, from list to list, and inter and intra list interferences could be studied. In addition, when *Ss* could not respond in the face of such liberal instructions, it was felt that the number and location of such blocks might also be revealing.

Procedure

In order to test the deductions from the unlearning hypothesis suggested above, it was decided to have *Ss* learn 10 successive lists at a rate and in a manner calculated to produce the greatest

proactive effects possible. In view of the consideration that unlearning was being investigated, it was desirable to overcome practice effects which might mask proactive effects, and, if the scores rose systematically, confuse the interpretation.

The hypothesis under test could not be studied adequately, unless each of the sets of syllables was of exactly equal difficulty. It was decided to avoid the problem of constructing sets to meet such an impossible criterion by counterbalancing the order of presentation in such a manner that each set of syllables occurred an equal number of times in each position. The development of such a counter-balancing arrangement proved to be feasible after considerable trial and error, and the solution of the 'Latin square' involved is presented as Table I. It will be noted that this method of arrangement not only meets the need of having each unit appear in every possible position, but also that each unit follows every other unit once and only once, thus avoiding any cumulative effects.[2] Such an arrangement should be used in learning single lists, the position of the various syllables being altered for different *S*s even when syllables of alleged equal association value are used because unequal association values are often created for adjacent syllables (construction of two-syllable words, sound or 'clang' associations, etc., are common practices of *S*s).

TABLE I

COUNTERBALANCED ORDER OF LEARNING OF 10 LISTS BY 10 GROUPS OF *S*s*

Lists	Subject Groups									
A	1	2	3	4	5	6	7	8	9	10
B	2	4	6	8	10	1	3	5	7	9
C	3	6	9	1	4	7	10	2	5	8
D	4	8	1	5	9	2	6	10	3	7
E	5	10	4	9	3	8	2	7	1	6
F	6	1	7	2	8	3	9	4	10	5
G	7	3	10	6	2	9	5	1	8	4
H	8	5	2	10	7	4	1	9	6	3
I	9	7	5	3	1	10	8	6	4	2
J	10	9	8	7	6	5	4	3	2	1

* The table shows the order in which each group learned the lists. Group I started with list A, then B, etc. Group II started with B, then D, etc.

In order to produce the maximum proactive effect, the method developed in the previous study, and later used by Underwood (6), was followed: in each of the 10 sets of 10 paired associates, the stimulus syllables were the same, and the response syllables differed. The convenient procedure previously reported of having the learning material mimeographed on slips of paper, so folded that a test could be taken on the same slip by merely turning the folded slip, was used. In each case, the test items were called for in a different arrangement from that in which they were learned. As far as the individual *S* was concerned, then, each had 10 slips of paper folded in the middle, with 10 pairs of syllables on the top half, and below the fold, invisible to the *S*, the 10 stimulus syllables, in a different order with blank spaces for the *S* to use in writing down his responses. The slips for each *S* were numbered from 1 to 10 and learned in that order. A letter from A to J identified the list.

Subjects.—One hundred college students in a beginning psychology course, unfamiliar with nonsense materials, were used. They were instructed in the mechanics of the experiment by preliminary illustrations with number-letter combinations and with word-word combinations. No hint of the purpose was given and the experiment was presented as a routine class room exercise in demonstrating learning procedures. The *S*s were instructed to learn as many as possible of the syllables on each slip with the implication that their ability to do well on the task would increase their understanding of the psychology of learning. It was decided to counteract possible

[2] Similar squares can be developed for any *even* number of items by following the pattern implicit in the square for 10 items. The basic procedure is to fill in the squares for each column by inserting numbers for pairs made up from opposite ends of the column, thus, 1 and 10, 2 and 9, 3 and 8, and so on. The totals for any pair should always equal the number of squares in the column plus 1.

general discouragement by announcing that it was expected that some confusion might be encountered, but to carry on in spite of any such effect and try to do as well as possible on each slip. The *Ss* were divided into 10 groups of 10 each. This arrangement was unknown to the *Ss*. No attempt was made to equate the 10 groups for learning ability as it was not felt that this was an important factor. Each group studied each of the 10 sets of syllables in the counterbalanced arrangement described above.

Instructions.—After preliminary description of the mechanics of the experiment, the following instructions were read to the *Ss*:

"Try to learn each list as well as possible. During the test periods write the correct response for each syllable as you come to it. If you do not know the correct answer, but some other answer occurs to you, write that down. If you later think of the correct answer, circle your wrong answer and write the correct one alongside. We are interested in the mistakes that you make just as much as in the correct answers. We would rather have you make a mistake than leave a blank. Do not leave blanks unless absolutely nothing seems to occur to you."

The device of circling wrong responses was introduced to encourage *Ss* to write down recognized errors without disturbing the normal set to obtain high scores.

Experimental routine.—To secure maximum proactive effects, the learning time for each set of syllables was limited to two min. A one min. period was allowed for each test (this appeared to prove sufficient in each case) and a 30 sec. pause followed each test. During the pause the *Ss* prepared themselves for the next learning session. No instructions were given following the start of the experiment.

Results

On the basis of the experimental procedure, data pertinent to an unlearning hypothesis and McGeoch's suggested explanation of unlearning as due to reproductive interference, are available in at least three forms: positive learning scores, blocks (no response), and remote associations. The data should be considered in terms of their bearing upon the following deductions:

1. When a strictly limited uniform time is allotted for learning each list, a second list will be poorly learned relative to a first, because, (a) some of the time assigned for learning the second list is wasted in unlearning the first, or (b) the well-learned first list interferes more effectively with the learning of the second; the consequently poorly learned second list cannot interfere as strongly with a third list, etc. This should result in an alternating improvement and decline in positive scores on successive lists with the odd lists better learned.

2. There should be relatively more blocks in the even-numbered lists because the better learned odd-numbered lists will either (a) intrude more effectively, or (b) take longer to unlearn, thus leaving less positive learning time.

3. If odd-numbered lists are freer from interference from preceding lists, they should have most of their interference from within their own lists (higher frequency of intra-list responses on the odd-numbered lists).

1. *Learning scores.*—The learning scores for each of the 10 learning series appear in Table II. It is immediately apparent that there is, in general, a significant decline from the first to the last list. This net loss corresponds with the expectations of a strong proactive effect

TABLE II

NUMBER OF CORRECT RESPONSES, BLOCKS, REMOTE (IDENTIFIED) ASSOCIATIONS, AND MISCELLANEOUS ERRORS PER TRIAL

List	Correct		Blocks		Misc. Errors Mean	Remote Ass'n Mean	Critical Ratios between Lists (No. Correct)	
	Mean	S.D.	Mean	S.D.			Between	C.R.
1	3.58	2.18	2.08	2.04	3.37	.99	1 and 4	2.17
2	3.31	2.14	2.16	2.01	3.43	1.24	1 and 6	3.67
3	3.39	2.12	1.94	1.85	3.21	1.46	1 and 7	3.34
4	2.95	2.05	2.05	2.00	3.39	1.81	1 and 8	3.65
5	2.95	2.02	1.94	1.86	3.49	1.86	1 and 9	3.94
6	2.55	1.71	2.05	1.98	3.46	2.18	1 and 10	3.59
7	2.61	1.97	1.88	2.13	3.66	2.17	3 and 6	3.11
8	2.52	1.94	1.63	1.97	3.90	2.07	4 and 9	2.11
9	2.36	1.96	1.65	2.24	4.15	1.89	5 and 6	1.53
10	2.47	2.20	1.68	2.05	3.99	1.93		
Total responses	2869		1926		3605	1760	Grand total: 10,140*	

* The total exceeds 10,000 because some *S*s added correct responses after making errors.

from the frustrating task imposed on the *S*s, both in terms of the conflicts developed in the use of the same stimulus syllables throughout the experiment, and in the short time allowed for learning each set. The manner in which the progressively low scores develop, however, is the chief factor of interest. It will be recalled that the hypothesis called for, in case of a general decrease in learning scores, a smaller loss on lists 3, 5, 7, and 9, and a relatively greater loss on lists 2, 4, 6, 8, and 10. It will be noted that such a relationship holds with striking regularity at least through series 8. The deviations of lists 9 and 10 from the hypothetical expectations will be considered later. The data have been plotted in Fig. 1a for easier inspection and comparison. It will be noted that, although a decrease was expected on each of the lists, list 3 shows an *increase*. List 5 shows no decrease whatsoever, remaining exactly on a level with list 4. List 7 again shows an increase. Lists 2, 4, 6, and 8, on the other hand, show relatively marked drops in learning scores. In general, then, it appears that the scores correspond with the experimental hypothesis in a manner that is better than was expected, as the hypothesis called only for a slighter degree of loss on odd-numbered series. The reliability of the points on the curve is difficult to establish. To calculate measures of significant differences is of no value in establishing the truth of the hypothesis, when, as in the case of series 5, there is *no* difference in the score from that of series 4, and the absence of the difference is the essence of the proof of the hypothesis; no difference in that case is what was expected. Where differences exist, critical ratios were calculated and appear in

Table II. Although the point-to-point differences are not reliable, there is a general reliability to the whole curve as it develops from points 1 and 3. In view of the fact that 100 *Ss* were used and that the standard deviations are reasonably uniform (although large), it is believed that whatever reliance can be placed on the variations in the

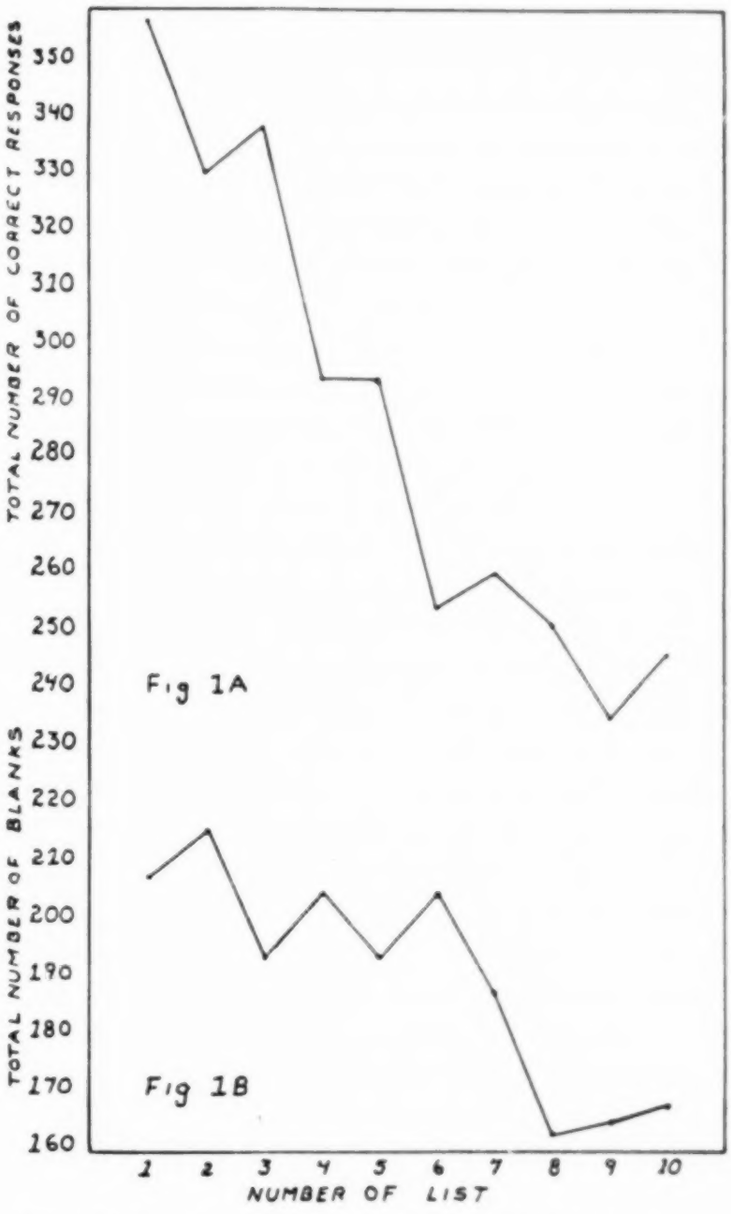


FIG. 1A. Learning scores per trial; FIG. 1B. Blocks per trial

curve must be based on the improbability of just such variations occuring by chance throughout eight successive points. Assuming that the probability of any point rising or falling is 0.5, then the probability of four successive falls and three intervening rises or non-falls would be 1 in 128. Leaving aside internal reliability, the conformity of the present results to those obtained previously by the present

writer and some of the results of Underwood suggests that the present results represent a genuine phenomenon.

The deviations from the trend shown by series 9 and 10 merit some comment. It is commonly noted in fatigue studies and other prolonged periods involving continuous exertion that an end-spurt occurs when a work session approaches a known or anticipated end. Such an end-spurt effect may be sufficient explanation of the rise in series 10; but observations of students in other experiments involving the learning of conflicting materials suggest another type of explanation which not only accounts for an end-spurt, but also for some of the preceding low performances. *Ss* have been observed to limit their performance voluntarily in such conflict learning situations when they realized that a relative mastery of current material might hamper performance on material to be learned in the future, when their experience has shown them that such effects do occur. When the *Ss* thus limit their performance, no true measure of capacity is obtained. As the task approaches its end, as for example, the tenth series in the present experiment is reached, the *Ss* no longer fear that learning the material will have any deleterious future effects and do exert themselves beyond their previous limited standards. Casual interrogation several of the *Ss* confirmed this suggestion, although no systematic attempt to study it has been attempted. The possibility of such an effect pervading the present experiment, or any similar experiment involving series of tasks, must be considered whenever *Ss* know the general outline of the procedure or the number of items to be used.

2. *Blocks in the learning process.*—In Table II are listed the number of times the *Ss* failed to respond in any way in each recall period. It will be noted from Fig. 1b that the number of failures to respond in any fashion (blank spaces left on the test sheet), follows almost exactly a reverse pattern from that of the learning scores, so that when learning scores are relatively high, the number of blank spaces is low, and vice versa. Considering the general unreliability of the separate points on the curve, it is possible, of course, that these failures to respond are artifacts of the situation; that, obviously, since less was learned on the even-numbered lists, there was more opportunity for failures to respond. On the other hand, just such a distribution was expected before the experiment was performed and *Ss* were specifically asked to try to avoid blanks with the open invitation to make errors instead. In view of the consistency of this trend, it is possible that the results merit more consideration than a mere dismissal. To expect such a pattern to occur by chance in light of the fact that there was ample opportunity to make some other response (correct responses and failures account for less than half of the 10 possible responses) seems unwarranted. Again assuming that by chance there

was equal probability of any point rising or falling, such a pattern would occur only once in 128 times. The probability, on such a basis, for both learning scores and failures to follow the pattern they did, is less than once in a thousand. The meaning of the pattern of increases and decreases in failures to respond is open to interpretation, either according to McGeoch's hypothesis or Melton's, and will be considered in the discussion.

3. *Remote associations.*—Each test sheet was carefully examined for errors which were exact (three letters) duplicates of syllables that had previously been learned on former lists and now 'intruded' into a list where they did not belong. Such intrusions or remote associations were then identified according to their origin by checking through previous lists to find where they did, in fact, belong. A

TABLE III

NUMBER AND PERCENT OF INTRA-LIST AND INTER-LIST INTRUSION ERRORS IN THE LEARNING OF EACH OF THE 10 LISTS, AND THE PERCENTAGE OF INTRUSIONS FROM EACH LIST ON THE SUCCEEDING LIST

Learned Lists	Lists from Which Errors were Made										Total Errors	Percent Intra-List Intrusions	Percent 1st Order Intrusions	
	10	9	8	7	6	5	4	3	2	1			List-List	%
1										99	99	100	1 to 2	.41
2									73	51	124	58	2 to 3	.26
3								89	39	18	146	61	3 to 4	.22
4							85	40	32	24	181	47	4 to 5	.18
5						75	33	28	27	23	186	41	5 to 6	.16
6					72	36	37	29	22	22	218	33	6 to 7	.14
7				69	32	20	24	22	21	29	217	32	7 to 8	.15
8			50	32	31	17	22	15	16	24	207	24	8 to 9	.14
9		57	26	15	18	9	25	15	10	12	189	30	9 to 10	.13
10	46	26	22	19	16	12	10	10	15	17	193	24		
Total errors from each list	46	83	98	135	169	169	236	250	255	319	1760			

record was also kept of all responses which came from the list under study, but were written down opposite the wrong stimulus. Such responses were considered intra-list errors or remote associations of zero-degree remoteness as far as the several lists were concerned. The distribution of such inter- and intra-list errors appears as Table III. It will be noted that they comprised approximately 17 percent of the total possible responses, the inter-list associations accounting for 10 percent by themselves. A number of suggestive relationships in the table deserve comment. It is at once apparent that responses learned in the very first list persist for some *Ss* all the way

through the ten series of learning tests. Here is direct evidence of interference with the learning process by intrusion of specifically identified responses. The influence of such errors remains fairly strong for several lists before the frequency of errors from a given list begins to wane and give way to its successors. In general, the greatest amount of the interference comes from the immediately preceding lists; this tendency is most marked with the early lists. By the time the *Ss* reached the fourth list, errors from lists 1 and 2 were more common than from list 3. This trend continued so that by the time that the tenth list was studied, the great majority of the associations (76 percent) were of more than one degree of remoteness. Because of the cumulative nature of the results, it is difficult to trace specific list-to-list relations in terms of the hypothesis under study. Another relationship which seems to hold in a general way should be considered. On the basis of an interference hypothesis and the demonstrated

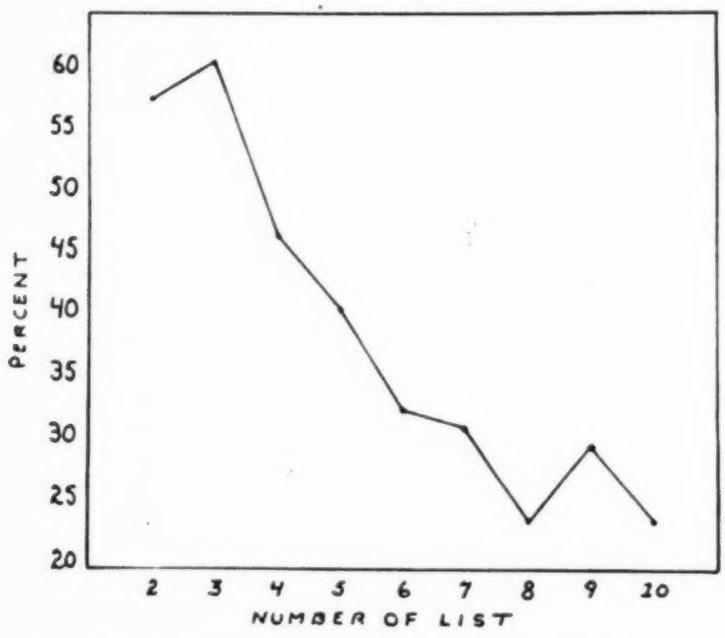


FIG. 2. Percent of intra-list intrusions per trial

poorer learning of the even-numbered lists, it can be assumed that odd-numbered lists should be relatively free from conflict from outside sources, because such interfering responses should have acquired only minimum strength in the poorly learned even-numbered lists. If, then, interfering responses could not be expected in great frequency from even-numbered lists, more opportunity would be available for confusion within the odd-numbered list itself, in addition to its being relatively better learned. To determine the possibility of such a relationship holding, the percent of intra-list associations was calculated on the basis of the total number of intra- and inter-list associa-

tions and plotted in Fig. 2. The value for list 1 is, of course, 100 percent, since there were no previous lists to interfere, and so this value is not shown. Starting with list 2, however, we see that its percent of intra-list associations is low compared with that of list 3; the decrease from 4 to 5 is not as steep as that from 3 to 4; 7 is practically the same as 6, and 9 is higher than 8 or 10. Thus, in every case the pattern holds good for the interference hypothesis.

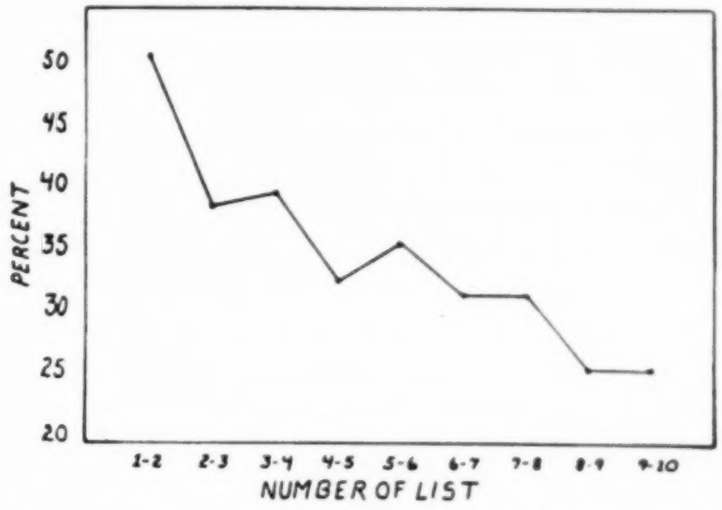


FIG. 3. Percent of intrusions from immediately preceding lists

Another interesting trend in the data again demonstrates the relative freedom of odd-numbered lists and the contrasting interfering effects from the odds on the evens. It is seen from Table III that 51 of the responses interfering in Test 2 came from list 1; only 39 interfering responses came from list 2 when Test 3 was taken, but 40 list 3 responses appeared on Test 4. The remaining odd-numbered lists always interfered more than did the even-numbered (5 on 6, 36; 6 on 7, 32; 7 on 8, 32; 8 on 9, 26; 9 on 10, 26). This relationship appears in Fig. 3 and again demonstrates the same principles.

DISCUSSION

The results obtained, in so far as their reliability can be assumed, correspond in general with an unlearning hypothesis even better than was expected. The trend of learning scores suggests that material which is poorly learned can be unlearned relatively easily, so that subsequent new material can be learned better. Such better learning, in turn, makes other future material harder to learn because it then requires more of any fixed time interval for its unlearning. The inability to demonstrate the reliability of the results may leave this interpretation still suspect in spite of numerous systematic trends demonstrated in

its support. Presumably use of sophisticated *Ss* equated for learning ability might reduce the variability of the scores. How this would help in demonstrating the significance of differences between scores when the absence of a difference is the significant finding, it is difficult to decide. In fact, the *E* is faced with a delimma when he attempts to refine his results by the suggested method; should he attempt such a refinement, it is likely that sophisticated *Ss* will learn with relative regularity (this is what possibly happened with Underwood's *Ss*). It is suggested that only repeated studies of this problem will eventually amass the data on which the hypothesis must rest.

Thus far, the term 'unlearning' has been used merely as a convenience for presenting the problem. It has been pointed out regularly in the presentation of the results that an interference explanation of the findings was equally logical even on the superficial grounds that a well-learned series would interfere with a subsequent series; this new series would then, being poorly learned, be unable to interfere effectively with the next series and so on. In McGeoch's interpretation of the problem, unlearning amounts to interference with interpolated learning by responses from an original list with consequent conflict and perhaps extinction of responses to the original list through failure of reinforcement. The data in the present experiment lend themselves strongly to such an interpretation, as this interference hypothesis is able to encompass all of the results while a simple unlearning hypothesis is able to encompass only some of them. The following generalizations from the data, while not necessarily in conflict with an unlearning hypothesis, are of no special use to it and do support a reproductive inhibition interpretation: (1) specific responses learned early continue to have an inhibitory effect on subsequent series; (2) as learning of successive lists continues, previously learned responses appear more and more frequently as interfering associations than do intra-list associations, the ratio of intra-list associations to total interfering responses drops from 58 percent in list 2, to 24 percent in list 10; (3) in addition to the directly identifiable remote associations and identifiable intra-list associations which reduce learning scores, 35 percent of the total response possibilities were made up of a variety of miscellaneous errors, many of which were recognizable as either misspellings of a previously learned response, an interference from a stimulus syllable, partial spellings of former responses, etc. The negative basis for the original unlearning hypothesis was the absence of sufficient intrusive errors. It appears from the willingness of the present *Ss* to make intrusive errors that all that is required to obtain a large number of such errors is to ask for them, and make it experimentally permissive to combat the normal attitudes of college

students to avoid obvious errors through the simple expedient of remaining silent.

A final comment should perhaps be added in support of the interference hypothesis. The writer has had occasion to use some *Ss* in nonsense syllable learning who learned similar lists of syllables as far back as nine months previously. These *Ss* reported interference effects in learning new lists from the old lists which they had presumably forgotten. Such reports have been common though unsolicited and the *Ss* report making strong efforts to get rid of these old associations by various stratagems which amount to voluntary measures aimed at discrimination, not elimination, nor unlearning. The cumulative frustration demonstrated in the progressively poorer learning and reported by the *Ss* in terms of confusion, strain, and effort, again support the interference hypothesis. Simple unlearning should not of itself create such instances of almost intolerable frustration.

Summary and Conclusions

1. One hundred *Ss* learned 10 lists of paired-associate nonsense syllables in rapid succession. The lists were designed to provide a high degree of interference.

2. The odd-numbered lists were learned relatively better than the even-numbered lists. On the assumption of an unlearning hypothesis this is just what should happen, since time spent in unlearning a former list would subtract from learning time on the list being studied. The results are equally susceptible of an interference explanation since a well learned list could interfere more strongly than one poorly learned.

3. *Ss* continued to respond with syllables learned in the early lists all through the later lists. Responses from list 1 were being made in list 10. There was a progressive decline in the appearance of such early learned responses, although intra-list interferences also tended to decline so that the majority of interferences were relatively remote.

4. McGeoch concluded his account of unlearning by leaving it to experiment to decide "whether this unlearning is different from the competition producing reproductive inhibition." The results of this experiment, on the whole, support McGeoch's interpretation of unlearning as reproductive inhibition at the time of interpolated learning.

(Manuscript received October 24, 1947)

References

1. Bugelski, B. R. Interference with recall of original responses after learning new responses to old stimuli. *J. exp. Psychol.*, 1942, **30**, 368–379.
2. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
3. McGeoch, J. A. *The psychology of human learning*. N. Y.: Longmans, Green, 1942.
4. Melton, A. W., & Von Lackum, W. J. Retroactive and proactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.
5. Twining, P. E. The relative importance of intervening activity and lapse of time in the production of forgetting. *J. exp. Psychol.*, 1940, **26**, 483–501.
6. Underwood, B. J. Associative inhibition in the learning of successive paired associate lists. *J. exp. Psychol.*, 1944, **34**, 127–135.

AN EXPERIMENTAL ANALYSIS OF SET IN ROTE LEARNING: THE INTERACTION OF LEARNING INSTRUCTION AND RETENTION PERFORMANCE

BY LEO POSTMAN AND WILLIAM O. JENKINS

Harvard University

Statement of the Problem

Set has often been considered a significant determinant of performance in a learning situation. For example, intent to learn is critical. Learning efficiency is greatly impaired if there is neither explicit nor self-induced instruction to learn (**3, 4**). The use of the term, *set*, has not, however, been limited to *S*'s intention to learn. At one time or another, *set* has been applied to many aspects of the learner's attitude (**1**). *Set*, then, is not a uniquely defined concept, but rather has many dimensions of variation. In view of the demonstrated importance of *set* as a determinant of learning, it is essential to specify these dimensions of variation and to analyze their operation experimentally.

One of the dimensions along which set may vary is *S*'s readiness for a particular type of retention test. During the period of practice, *S* knows or infers from *E*'s instructions under what conditions the results of training will be tested. In many cases, as for example in the method of anticipation, practice and tests proceed concurrently. The types of associations formed among the learning materials will necessarily reflect the test procedure by which they are to be measured. The learner may emphasize different relationships among the stimuli and responses and may rehearse different ways of producing the responses.

The purpose of these experiments was to show how readiness for a particular type of test influences learning performance. The specific hypothesis was that performance would be optimal if the retention test was the one expected by *S* and that performance would be impaired if the test was other than the one expected by *S*.

Experimental Procedure

It was necessary to hold the learning materials and the amount of practice constant and to vary only the relationship between instruction and test.

The learning materials consisted of 25 two-syllable adjectives. *E* read these items to a given

group of *Ss* at the rate of one every two sec. The list was read through five times, with an interval of six sec. between successive presentations of the list.

Three sets of learning instructions were used: instruction for anticipation, free recall, and recognition. Correspondingly, three test procedures were employed: anticipation, free recall, and recognition. Each instruction was paired with each one of the three kinds of test, yielding nine instruction-test combinations.

Since learning and test instructions constituted the critical variables in this experiment, they are presented in detail below.

Learning Instructions

Anticipation.—This is an experiment in learning. I am going to read to you one by one a list of 25 two-syllable words. I want you to listen to them carefully. I am going to read the list through several times. Then your memory for these words will be tested. You will be asked to anticipate the words before I say them. When I say 'Go', you will write down the first word in the list before I have spoken it. When I have read the first word of the list, you will write the second word of the list, and so on. For example, if the first two words on the list were *table-window*, when I say 'Go' you would write *table*; when I say *table*, you write *window*, and so on.

Free recall.—This is an experiment in learning. I am going to read to you, one by one, a list of 25 two-syllable words. I want you to listen carefully. I am going to read the list through several times. Then your memory for these words will be tested. You will be given a sheet of paper on which to write down all the words that you can remember. You will be given a certain amount of time in which to write down all the words in the list that you remember.

Recognition.—This is an experiment in learning. I am going to read to you one by one a list of 25 two-syllable words. I want you to listen carefully. I am going to read the list through several times. Then your memory for these words will be tested. You will be given a long list which will contain the words I have read mixed in with a large number of other words. It will be your task to check the words which you recognize as being the words that I called out.

Test Instructions

Anticipation.—Now I want you to anticipate the words before I read them. When I say 'Go', write the first word in the list in the space marked No. 1. Write it down *before* I have said it. When I read the first word, write the second word in the space marked No. 2, and so on. Always write down the word before I have read it. In no case write a word down after I have read it or correct what you have written. If you have started writing a word when I begin to read it, you may finish it, but exactly the way you intended to write it. As soon as you have written a word, cover it with the strip of cardboard. If you cannot anticipate a word, put a line through the appropriate space. The important thing is to anticipate the words before they are read.

Free recall.—On this sheet of paper write down all the words that you can remember. If possible, write them in their right order in the list—notice that your sheet is numbered 1 through 25. If you cannot place a word in its proper position, but just remember that it was somewhere in the list, write it down anyway. Work rapidly since you will have only a limited amount of time. Start when I say 'Go'.

Recognition.—On this sheet of paper you will find 100 words. Among these 100 words are the words which I have read to you before. I want you to underline those words which you recognize as having been read to you. Go through the list rapidly once and do not linger over any one word. Turn the paper over as soon as you have gone through it once. Start when I say 'Go'.

The test instructions were prefaced by the following statement, if the instructions were changed from the original ones:

The memory test will be different from the one originally announced. Try your best, however, to remember as much as you can.

The test was given immediately after the fifth presentation of the word list. When the anticipation method was used, there was one test trial, four sec. being allowed for the anticipation of each successive item. For the free recall test, a time limit of 2.5 min. was set. In the method of recognition, a test sheet of 100 alphabetized adjectives was used containing the original 25 items. *Ss* were allowed a maximum of three min. for this test. Preliminary experimentation had shown that these time limits would yield fairly complete retention scores.

Ss were men and women students in an undergraduate course in experimental psychology. The experiment was conducted as part of the regular classroom work. Until the completion of the experiment, however, *Ss* were kept in ignorance of its purpose. Groups of *Ss* were assigned to the different experimental conditions at random. A total of 215 *Ss* participated.

Results and Discussion

Two sets of scores were obtained for the anticipation tests and the free recall tests: (1) number of items correct without regard to position in the series, and (2) number of correct items given in their proper positions. The results of the recognition tests, of course, required a different scoring procedure. In this latter case, the scores were obtained by subtracting the number of incorrect recognitions from the number of correct ones.

Retention Scores

The mean number of items correctly retained without respect to position, under all conditions of the experiment, appears in Table I, along with the number of cases on which each of these averages is based.[1]

TABLE I

Average Number of Items Retained Under the Different Experimental Conditions
Number of cases on which means are based appear in parentheses.

Test Procedure	Learning Instruction		
	Anticipation	Free Recall	Recognition
Anticipation	9.90 (40)	7.96 (25)	8.25 (20)
Free recall	15.00 (21)	13.60 (25)	11.17 (24)
Recognition	18.58 (12)	19.26 (23)	21.20 (25)

There are three systematic sources of variation: the learning instructions, the testing procedure, and the interaction of these two. To evaluate the significance of the contribution made by each of these factors an analysis of variance was performed. In Table II are the

[1] It will be noted that *N*s are fairly homogeneous in seven of the cells (*ca.* 25), but deviate considerably in two cases (12 and 40 respectively). This fact was, of course, taken into account in the statistical analysis.

TABLE II

ANALYSIS OF VARIANCE ON RETENTION SCORES

	df	*F*	*P*
Instructions	2	1.45	>.05
Tests	2	127.69	<.01
Interaction	4	3.69	<.01

results of the *F*-tests, with an error variance of 16.66 based on 206 degrees of freedom:

Instructions, in and of themselves, do not make a significant difference. Tests *per se*, on the other hand, contribute greatly to the differences in retention scores. This is well in accord with previous findings (3). Recognition yields the highest degree of retention, anticipation the least, with free recall occupying an intermediate position. It is the *interaction of learning instructions and testing precedures* which provides the critical datum in these findings. The reliable interaction effect indicates that retention varies significantly with the relationship between instruction and testing procedure.

To gauge the nature of this interaction, it is necessary to proceed to a detailed examination of the differences between the various groups. For this purpose, we compare, for each of the test procedures the results obtained with different instructions.[2]

Anticipation tests.—Both free recall and recognition instructions lead to poorer retention than instruction to anticipate (*t*-ratios of 1.94 and 1.50 with *P*-values of .03 and .07, respectively).

Free recall tests.—*S*s set for a recognition test do worse that *S*s ready for free recall ($t = 2.0$ and $P = .02$). It is interesting to note that instruction to anticipate not only fails to affect the free recall test adversely, but even works slightly in the opposite direction: those instructed for anticipation do somewhat better on the free recall test ($t = 1.17$ and $P = .23$).

Recognition tests.—The performance of those instructed for recognition is reliably better than that of those ready for an anticipation test ($t = 1.85$ and $P = .03$) or a free recall test ($t = 1.58$ and $P = .06$).

What, then, is the nature of the interaction between instructions and retention tests as it manifests itself in differences between specific experimental groups? Clearly it is not sufficient to say that sheer similarity between learning instruction and retention problems leads to optimal performance. This hypothesis would account for two cases —anticipation tests and recognition tests—but fails to explain the findings on the free recall tests. When retention is tested by free recall, instruction for anticipation does not interfere with performance

[2] In carrying out the individual *t*-tests an improved estimate of variance is based on the error variance of the total distribution divided by the appropriate number of degrees of freedom (2, p. 96 f.). The value of *t* is then evaluated in terms of one tail of the distribution.

but even facilitates it slightly. This relationship is, moreover, asymmetrical since instruction for free recall impairs performance on the anticipation test.

What is needed, then, is a more general hypothesis which will reasonably account for the total interaction effect. The following generalization suggests itself: to the extent that the learning instructions single out and emphasize those aspects of the learning task which will be used in the test situation, performance is favored; to the extent that learning instructions fail to establish a set toward the aspects of the material most relevant to the test, retention is impaired. This generalization accounts for those cases in which sheer similarity of instruction and test seems to favor performance as well as for the case which at first glance seemed to offer a contradiction to the general trend of results. Instruction to anticipate gets *S* ready actively to reproduce as many items in the series as he can. To that extent it overlaps with the instruction for free recall. In addition, expectation of an anticipation test focuses *S*'s attention on serial position and utilization of intraserial cues. Anticipation thus is similar to free recall but more exacting. The fact that instruction for anticipation does not hinder performance on a free recall test but even tends to favor it is thus in accord with our generalization, for instruction for anticipation emphasizes aspects of the learning task which are relevant and useful for a free recall test, in addition to those aspects which are peculiar to anticipation. On the other hand, instruction for free recall fails to emphasize what is critical to anticipation—serial position—and therefore impairs performance on the anticipation test. The asymmetry of the realtionship may thus be ascribed to the greater rigor and difficulty of the anticipation procedure.

Position Scores

The average number of items reproduced in their correct positions in the anticipation and free recall tests is presented in Table III.

TABLE III

AVERAGE NUMBER OF ITEMS REPRODUCED IN THE CORRECT SERIAL POSITION UNDER THE DIFFERENT EXPERIMENTAL CONDITIONS

Number of cases on which means are based appears in parentheses.

Test Procedure	Learning Instruction		
	Anticipation	Free Recall	Recognition
Anticipation	8.05 (40)	6.30 (25)	6.64 (20)
Free recall	6.43 (21)	5.16 (25)	2.25 (24)

These results closely parallel the retention scores obtained without regard to position (cf. Table I). The same sources of variation must be considered: learning instructions, tests, and the interaction of these two. In this case, all three factors make a significant contribution. In Table IV are the results of the *F*-tests, with an error variance of 15.01 based on 149 degrees of freedom:

TABLE IV

ANALYSIS OF VARIANCE UPON POSITION SCORES

	df	*F*	*P*
Instructions	2	8.41	<.01
Tests	1	18.38	<.01
Interaction	2	23.33	<.01

It is not surprising that instruction in and of itself becomes a significant variable when position scores are considered. To retain a series of items *in their proper order* is a special task imposed on *S* in addition to the general directions to remember. *S*s will presumably neglect this aspect of the learning task without explicit instruction. As for the two tests, the anticipation method yields the higher number of correct position responses. This finding may be ascribed to the position cues provided by this testing procedure. The most important datum is again the significant interaction of instructions and tests. Comparison of the individual groups yields results highly consistent with those found for the sheer retention scores. Again test performance is impaired if instruction and retention problem do not harmonize in their selective emphasis on certain aspects of the learning task.

SUMMARY

The results of this experiment reaffirm the importance of set as a basic determinant of learning performance. Systematic manipulation of the learner's set, with stimulus materials and amount of practice held constant, yields highly significant results.

The learning materials were 25 two-syllable adjectives. The list was presented to all *S*s five times at a constant rate. Set was varied by instructing the subjects for different types of retention tests: anticipation, free recall, and recognition. Retention was tested by these three methods. All possible combinations of tests and instructions were used so that every instruction was followed by every test. To the extent that learning instructions single out and emphasize those aspects of the learning task which are to be used in the retention test, *S*s' performance is favored. To the extent that learning instruction and retention problem fail to harmonize in their selective emphasis on certain aspects of the learning task, performance is impaired.

These findings emphasize the need to isolate the specific dimensions along which set may vary. Clearly one of these dimensions is *S*'s readiness for a particular type of test. Throughout the period of practice, the expectation of a specific testing procedure influences learning behavior.

(Manuscript received January 10, 1948)

References

1. Gibson, J. J. A critical review of the concept of *set* in contemporary experimental psychology. *Psychol. Bull.*, 1941, **38**, 781–817.
2. Lindquist, E. F. *Statistical analysis in educational research.* Boston: Houghton-Mifflin, 1940.
3. McGeoch, J. A. *The psychology of human learning.* New York: Longmans Green & Co., 1942.
4. Postman, L., & Senders, V. L. Incidental learning and generality of set. *J. exp. Psychol.*, 1946, **36**, 153–165.

INTENSITY OF THE CONDITIONED STIMULUS AND STRENGTH OF CONDITIONING: I. THE CONDITIONED EYELID RESPONSE TO LIGHT[1]

BY DAVID A. GRANT AND DOROTHY E. SCHNEIDER

University of Wisconsin

Introduction

The present experiment was designed to reveal the effect of intensity of the conditioned stimulus (CS) upon the strength of conditioning. The human eyelid response was conditioned using a corneal air-puff as the unconditioned stimulus and light as the CS.

The relationship between CS intensity and strength of conditioning is an important one despite the lack of evidence as to its nature. It is one of the basic relationships between a fundamental stimulus variable and the fundamental dependent variable of learning. Knowledge of the function relating these two variables should lead to a better understanding of the process of conditioning. Exploration of this relationship constitutes also a direct attack on the problem of the influence of magnitude of the *stimulus trace* on the learning process (**12**, Ch. XI).

The objectives of this investigation necessarily pose an interesting problem in experimental design, a problem which apparently has impeded prior efforts. The independent variable, CS intensity, and the dependent variable, strength of conditioning or strength of the *functional linkage* between CS and CR, are clearly defined. However, a number of investigators have reported that, with a given level or strength of conditioning, response magnitude varies with CS intensity (**16**, **17**, **18**, **20**). As pointed out by Hilgard and Marquis (**8**, pp. 141–144), the straight-forward approach of conditioning two or more groups of *S*s with different levels of CS intensity *confounds strength* of *conditioning* with *response strength* and is, therefore, inadequate. An alternate approach would be to reinforce with two or more CS intensities and extinguish with a single common intensity. As resistance to extinction is a good measure of strength of classical conditioning (**11**, **8**), this approach is promising; but a difficulty arises because the extinguished responses would be *generalized* CRs which extinguish more rapidly than the original CRs (**10**).

[1] This research was supported in part by the Research Committee of the Graduate School of the University of Wisconsin from special funds voted by the State Legislature for 1946–47.

A satisfactory factorial design (**22**) must make strength of conditioning completely orthogonal to or independent of response strength. In the present study this was achieved by using 16 experimental groups. These groups may be thought of as forming a square of four rows and four columns. The four rows differed according to the *intensity of the CS during reinforcement*, and the four columns differed according to the *intensity of the CS during extinction*. The experiment thus provided for the combination of all four CS intensities during extinction with all four CS intensities during reinforcement. When extinction measures are set into the resulting 16-fold table, row differences would be due to the influence of CS intensity upon strength of conditioning, and column differences would be due to the influence of CS intensity upon response strength. In this way, the two dependent variables are isolated. If significant *interaction* appears, it can be attributed primarily to generalization effects.

Apparatus and Procedure

The apparatus and general procedure were similar to those of previous experiments (**7, 5**) in which a photochronograph was used to present light as the CS and a corneal air-puff as the UCS. Light adaptation of the *S*s was preserved throughout the experiment to prevent confusion of the β-response with the conditioning results (**3, 4**). The milk-glass window from which the CS was presented was maintained at 27 millilamberts brightness after a brief initial exposure of over 1500 millilamberts designed to bring all *S*s to a uniform state of light adaptation. The four light stimuli consisted of raising the 27 millilamberts pre-stimulus brightness by an additional 7, 70, 320, or 1050 millilamberts for a period of .75 sec. The sequence of trials was as follows: Day 1, 25 *Reinforcement trials* with *reinforced* CS and puff; Day 2, 25 *Reinforcement trials* as on Day 1, 2-min. rest; 15 *Extinction trials* with *extinguishing* CS, 2-min. rest; 5 *Re-extinction* trials with *reinforced* CS.[2] The interval between trials was varied unsystematically from 30 to 50 sec.

The *S*s were 64 men and women students from elementary classes in psychology. Four *S*s were assigned (according to order of appearance) to each of the 16 experimental groups described in the *Introduction*, so that the 63 degrees of freedom were divided as indicated in the first two columns of Table II, below.

Results

As stated above, the experiment was conducted under light adapted conditions designed to avoid sensitizing the β-response to light. Under these conditions the CRs were restricted fairly well to the latency range from .25 to .50 sec. Eyelid responses occurred most frequently in this range and included all responses following typical CR acquisition and extinction trends. This observation, which is in line with other findings (**5, 6**), justified restriction of the analysis to the responses in this latency range.

The frequency and magnitude of the CRs increased regularly in negatively accelerated curves during reinforcement and decreased

[2] The addition of the re-extinction series, suggested by Dr. Margaret Kuenne, permitted a test of the intensive generalization of extinction.

rapidly in negatively decelerated curves during extinction. The regularity of these curves indicated that the experimental procedure resulted in typical eyelid conditioning.

1. *Strength of conditioning and response strength.*—The extinction data are summarized in Table I. The 16 cells of Table II are arranged in a four row by four column pattern. Each cell contains two entries: (1) mean frequency of CRs per *S* during the 15-trial

TABLE I

SUMMARY OF MEAN FREQUENCY AND MEAN TOTAL MAGNITUDE (MM.) OF CRS DURING EXTINCTION

	Stimuli	Conditioned Stimulus Intensity During Extinction (Millilamberts)					Row Marginals
			7	70	320	1050	
Conditioned Stimulus Intensity During Reinforcement (Millilamberts)	1050	Freq.	0	3.25	5.00	5.25	3.38
		Mag.	0	31.75	55.00	60.75	42.50
	320	Freq.	1.25	3.75	3.50	4.25	3.19
		Mag.	29.50	111.00	42.75	61.75	61.25
	70	Freq.	4.25	1.00	4.50	1.25	2.75
		Mag.	38.25	2.00	71.50	14.00	31.44
	7	Freq.	5.25	2.00	2.25	1.75	2.81
		Mag.	77.50	36.00	55.50	26.50	48.88
	Column marginals	Freq.	2.69	2.50	3.81	3.12	3.03
		Mag.	36.31	45.19	61.81	40.75	46.02

Extinction Series and (2) mean total magnitude of CRs during the Extinction Series. The frequency and magnitude scores were highly reliable. In both cases the odd-even reliability coefficient was 0.96. The row and column marginal means are given on the right-hand and bottom margins, with the general means in the lower right-hand corner.

Row marginal means for frequency and magnitude show little tendency to increase with increasing intensity of the CS during reinforcement. The trend is very irregular. *Column* marginal means become generally larger as a function of increased stimulus intensity, but the upward trend is also very irregular. The statistical significance of the row and column marginal variation was tested by means of analysis of variance. A summary of these analyses is presented in Table II in the two columns headed *Extinction.* Neither the row nor the column variation is statistically significant. The fact that the row variation is not significant means that strength of conditioning cannot be claimed to be a function of variation in intensity of the CS. The absence of statistical significance in the column variation means

that variation in stimulus intensity did not produce reliable variation in response strength.

2. *Generalization effects.*—As stated earlier, interaction between rows and columns in the present experimental design would be produced chiefly by intensive generalization effects. A rough generalization curve may be constructed by averaging diagonally the entries of Table I. These curves for both frequency and magnitude are approximately what might be expected for intensive generalization (9). There are progressively greater drops in frequency and magnitude of CRs with increasing differences between reinforcing and extinguishing stimuli. The interactions were not statistically significant, however, so that the generalization trends cannot be considered reliable.

TABLE II

MEAN-SQUARES SUMMARIZING THE ANALYSIS OF VARIANCE FOR EXTINCTION AND RE-EXTINCTION FREQUENCY AND MAGNITUDE DATA

Source of Variation	*df*	Extinction		Re-extinction Magnitude
		Frequency	Magnitude	
Reinforcement intensity (Rows)	3	1.44	2480.77	313.69
Extinction intensity (Cols.)	3	5.44	1984.56	607.19
Generalization (Interaction)	9	16.65	4517.70	660.43
Error variance (Replication)	48	13.51	5928.99	856.29

Scores analyzed are total frequency and total magnitude rather than means. None of the mean-squares is statistically significant.

After the initial 15-trial Extinction Series, each *S* received an additional 5-trial Re-extinction Series, returning to the same intensity of CS which had been used during reinforcement. This permitted a test of the generalization of extinction. The average frequency and magnitude of the CRs during the Re-extinction Series were most reduced when the extinction stimulus and re-extinction stimulus were identical and least reduced when they were widely different in intensity. These results were as expected, but the data were not statistically significant in spite of the fact that the individual scores were highly reliable (coefficients of 0.91 for frequency and 0.94 for magnitude). The magnitude data were subjected to analysis of variance, and the analysis is summarized in the right hand column of Table II. The interaction effect (as well as the row and column effects) is not statistically significant. Thus, there is no significant evidence for generalization of extinction.

Discussion

The chief finding of this experiment is that variation in CS intensity had no statistically significant effect upon the strength of conditioning; and the design employed appears to be the first systematic study capable of revealing this fact. Although there was a slight trend in the expected direction of stronger conditioning with greater intensity, the statistical tests yielded negative results. The negative finding is of considerable importance because of the high reliability of the extinction scores and because of the great power of the *F*-test based upon 48 *df* for the error variance and three *df* for the numerator (**2**). These negative results were presaged by Hilgard and Marquis's (**8**, p. 142) reanalysis of Hovland's GSR data and by Carter's findings (**1**) with the eyelid response to sound.

The absence of a clear-cut effect of variation in CS intensity upon *response strength*, as opposed to strength of conditioning, seems, at first glance, to be contrary to previous reports (**16**, **17**, **18**, **20**); but a close examination of the earlier studies reveals that their results were in no instance based upon a number of cases sufficient to give statistical reliability. The insignificant trends of the present study are thus actually in line with those of earlier studies.

The negative results must, of course, be interpreted in terms of the stimuli and response used. Had a greater range of stimuli been used, perhaps the results would have been positive. For example, stimuli near threshold might have had greater effects on response strength (**18**), although their effects upon resistance to extinction might have been complex (**13**, **14**, **15**). It was also thought that utilization of another response mechanism or another modality could lead to other results, but further work along these lines has been completed at the University of Wisconsin[3] only to confirm the present data. One further possibility is that the effects of CS intensity may not be revealed until after extended periods of reinforcement (**20**). This possibility should be explored more thoroughly than was permitted by the scope of the present study.

The findings have two important implications. First, the hypothesis that rate of learning or habit formation depends on the intensity of the stimulus trace (**12**, Ch. XI) must be considered as being only a crude approximation or, at most, a possibility; and, consequently, interpretations based upon this hypothesis (e.g., **21**) must be treated with caution. Finally, the results favor Pavlov's old interpretation (**19**) of the role of the CS as that of a signal which sets off a conditioned triggering mechanism in the organism. If this interpre-

[3] Intensity of the conditioned stimulus and strength of conditioning: II. The conditioned galvanic skin response to an auditory stimulus. (In press, this Journal.)

tation is correct, considerable reorientation of conditioning theory will be required. Attention will have to be focused on: (1) the process of initially setting up or constructing the triggering mechanism, and (2) the process of setting the trigger itself.

Summary

Strength of conditioning the eyelid response to light was investigated as a function of the intensity of the CS. Strength of conditioning was separated from response strength by means of a special factorial design. Sixteen experimental groups of four *S*s each were used. The 16 groups might be thought of as forming a four row by four column square in which the rows differed according to the four brightnesses of light stimuli which were used as conditioned stimuli during the reinforcement trials. The four columns differed according to the four CS intensities which were used during the extinction trials. The four stimulus brightnesses were 7, 60, 320, and 1050 millilamberts. Fifty reinforcement trials, consisting of paired presentations of light and corneal air-puff, were given, 25 trials on each of two days. Fifteen extinction trials were given after the reinforcement trials of day 2. Re-extinction Series were then given with the same CS that had been used during reinforcement. All trials were run with the *S*s light adapted. The frequency and magnitude of CRs during extinction were analyzed by means of analysis of variance, and the principal findings were as follows:

1. Although there were trends of increased frequency and magnitude of CRs with the higher intensities of the CS, neither strength of conditioning nor response strength were significantly affected by variation in CS intensity.

2. Generalization both of reinforcement and extinction seemed to be present, but these tendencies were not statistically significant.

3. It was pointed out that, considering the highly reliable frequency and magnitude scores and the great power of the *F*-test based upon 48 *df*, the negative findings should be given special weight, although the results might be specific to the eyelid response and the range of stimuli which were used.

4. Since strength of conditioning was not affected by CS intensity, it was suggested that rate of habit formation could scarcely be expected to be a function of the intensity of the stimulus trace as in Hull's formulation. Instead, an interpretation of conditioning in terms of signalling and triggering mechanisms would be favored.

(Manuscript received October 24, 1947)

References

1. Carter, L. F. Intensity of conditioned stimulus and rate of conditioning. *J. exp. Psychol.*, 1941, **28**, 481–490.
2. Ferris, C. D., Grubbs, F. E., & Weaver, C. L. Operating characteristics for the common statistical tests of singificance. *Ann. Math. Statist.*, 1946, **17**, 178–197.
3. Grant, D. A., A sensitized eyelid reaction related to the conditioned eyelid response. *J. exp. Psychol.*, 1945, **35**, 393–402.
4. Grant, D. A., & Norris, Eugenia B. Dark adaptation as a factor in the sensitization of the beta response of the eyelid to light. *J. exp. Psychol.*, 1946, **36**, 390–397.
5. Grant, D. A., & Norris, Eugenia B. Eyelid conditioning as influenced by the presence of sensitized beta-responses. *J. exp. Psychol.*, 1947, **37**, 423–433.
6. Grant, D. A., Hake, H. W., & Schneider, Dorothy E. Effects of pre-testing with the conditioned stimulus upon the extinction of the conditioned eyelid response. *Amer. J. Psychol.*, 1947, **60**, 243–246.
7. Hilgard, E. R., & Campbell, A. A. The course of acquisition and retention of conditioned eyelid responses in man. *J. exp. Psychol.*, 1936, **21**, 227–247.
8. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: Appleton-Century, 1940.
9. Hovland, C. I. The generalization of conditioned responses. II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51**, 279–291.
10. Hovland, C. I. The generalization of conditioned responses. III. Extinction, spontaneous recovery, and disinhibition of conditioned and of generalized responses. *J. exp. Psychol.*, 1937, **21**, 47–62.
11. Hull, C. L. Learning: II. The factor of the conditioned reflex. In: C. Murchison, edit., *A handbook of general experimental psychology.* Worcester, Mass: Clark Univ. Press, 1934.
12. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
13. Humphreys, L. G. The effect of random alternation of reinforcement upon the acquisition and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, **25**, 141–158.
14. Humphreys, L. G. Acquisition and extinction of verbal expectation in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, **25**, 294–301.
15. Humphreys, L. G. Extinction of galvanic skin responses following two conditions of reinforcement. *J. exp. Psychol.*, 1940, **27**, 71–75.
16. Kupalov, P. S., & Gantt, W. H. The relationship between the strength of the conditioned stimulus and the size of the resulting conditioned reflex. *Brain*, 1927, **50**, 44–52.
17. Kupalov, P. S., Lyman, R. S., & Lukov, B. N. The relationship between the intensity of tone-stimuli and the size of the resulting conditioned reflexes. *Brain*, 1931, **54**, 85–98.
18. Newhall, S. M., & Sears, R. R. Conditioning finger retraction to visual stimuli near the absolute threshold. *Comp. Psychol. Monogr.*, 1933, **9**, No. 43.
19. Pavlov, I. P. *Conditioned reflexes.*, London: Oxford Univ. Press, 1927.
20. Podkopaev, N. A. (Further contribution to the correlation between the values of the conditioned reflex and the unconditioned food reflex.) *Trud. fiziol.* Lab. Pavlov., 1945, **12**, 124–133. Cited from *Psychol. Abstr.*, 1946, **20**, 427.
21. Reynolds, B. The acquisition of a trace conditioned response as a function of the magnitude of the stimulus trace. *J. exp. Psychol.*, 1945, **35**, 15–30.
22. Snedecor, G. W. *Statistical methods.* Ames, Iowa: Collegiate Press, 1946.

PROACTIVE INHIBITION AS AN EFFECT OF HANDEDNESS IN MIRROR DRAWING

BY CHARLES W. SIMON *

Antioch College

Introduction

The present experiment grew out of an observation by the writer of two children who were playing with the mirror drawing apparatus. One child was right-handed; the other, left. The writer noticed that the right-handed child did better with her left hand, while the left-handed child did better with her right hand. A review of the literature on mirror drawing brought to light experimental evidence suggesting that other *Ss* also did better with the non-preferred hand.

P. H. Ewert (6) used the mirror drawing task to illustrate bilateral transfer of training. His study included data which made first trials of right and left hand performances comparable. Though he was not examining the present problem, his results indicated in five out of six comparisons—two between the same individual's scores and four between independent group averages—a tendency in favor of the non-preferred hand over the preferred hand. This effect was present for both time and errors in three comparisons, and for errors alone in the two others. C. W. Bray (2) wrote: "Ewert was able to show that transfer is greater from the preferred to the non-preferred hand than in the opposite direction."

T. W. Cook (4), in his studies of cross-education between hands and feet, listed certain tables that made average first trial performances between right and left hands comparable for two groups of 10 *Ss*, only one of whom was left-handed. Though Cook's data are complicated by records of 'foot performance,' he found a definite superiority of the right hand over the left when the mirror was not used, and a left hand superiority for the same *Ss* when the mirror was used. This was true, no matter which hand was used first.

Cook (5) attempted to explain this left hand superiority in terms of cross-education resulting from practice with the left foot, and further generalized that training in one set of muscle groups results in a superiority by

* The writer wishes to express his appreciation to Dr. Francis W. Irwin, University of Pennsylvania, without whose guidance and counsel this experiment would not have been made; to Miss Constance Simon and Miss Suzanne Booth for their continued readiness to act as subjects in this and other experiments; to the Boy Scouts' Troop 49 and the Girl Scouts' Troop 200 of Philadelphia, Pa., for providing subjects with whom to work; to Dr. Clarence Leuba, Antioch College, for his assistance in the preparation of this manuscript; and to the many friends who volunteered as subjects.

cross-education of asymmetrical sets of muscles on the same side of the body when compared to the opposite side of the body. It is equally probable, however, that these results obtained from the mirror tracing task actually gave evidence of a superiority in the non-dominant hand, for Cook failed to perform an experiment having practice trials with the right foot to see if the right hand would then become superior to the left one.

L. E. Travis (11) closely approached an investigation of this phenomenon in his study of right-handed stutterers and normal speakers. He found that right-handed stutterers were more facile with their left hand in a mirror drawing task, while right-handed normal speakers were more facile with their right one. Comparing right and left hand performances of all these right-handed *Ss*, however, it was found that the mean time taken to complete the star was less for the left hand, though the difference appears not to have been statistically significant. Travis used no left-handed *Ss*, and all of his *Ss* performed with their left hand first.

The above observations, as previously stated, suggest a possible superiority of the left hand for right-handed *Ss* in mirror tracing., In two of the above experiments, the handedness of the *Ss* was given though actual methods of determining this handedness were not always explicitly explained. Since only approximately six percent of the general population are left-handed (1), it is probably safe to assume, even in the third case, that the results as a whole were primarily from a right-handed population.

Three questions arise: is this apparent superiority of the left hand in the mirror drawing of right-handed *Ss* really a fact; is the converse true with left-handed individuals; and, if so, are these phenomena subject to any limitations which may prevent them from being exhibited in full, as in the above experiments?

The present experiment was devised to test the hypothesis that, subject to secondary effects, the non-dominant hand is superior to the dominant hand in the mirror drawing situation. This hypothesis is supported by the data described above and is in harmony with the principle of negative transfer, or proactive inhibition. It is likely that old sensori-motor connections built up between the eyes and the dominant hand are stronger than those between the eyes and the non-dominant hand. It would therefore be more difficult for the dominant hand to break down these connections and to establish the new ones necessary in this task. This would result in a superiority of the non-dominant hand in this situation. To test this hypothesis, the following experiment was carried out.

Procedure

Sixty volunteer *Ss*, chosen from several Boy and Girl Scout troops, from among *E*'s acquaintances, and from a college student body, were used in the experiment. Forty of the *Ss* were right-handed, and 20 were left-handed. Of the right-handed *Ss*, there were 23 males and 17

females, ranging from 12 years to 31 years of age, with a mean age of 20.4 years. Of the left-handed *S*s, there were 11 males and 9 females, ranging from 13 years to 24 years of age, with a mean age of 19.3 years. All *S*s were white, and none had had controlled mirror drawing experience before.

Before being presented with the mirror drawing task, each *S* was given a test for handedness. The criterion for manual dominance was that the *S* must perform at least five out of the seven tasks described below with the same hand that he used to write his name; only persons fulfilling this criterion were used.

These seven tasks for the handedness test were chosen from the 45 questions of the taught and untaught unimanual tests found in Helen L. Koch's (8) 105 item questionnaire. Only the unimanual tasks were considered, for the performance in mirror drawing was essentially unimanual. Also, these taught and untaught unimanual groups correletad higher (.97) with Koch's entire 105 item test and higher (approximately .94) between test-retest variations of performance-questionnaire combinations than any other group. These taught and untaught unimanual test groups were higher than any other group when correlated by the split-half method (.91 and .88 respectively), a fact which Koch felt showed generality and stability. From the 45 tests in these two groups, using Koch's statistics, the seven individual tests chosen and adapted to this experiment were those which correlated .70 or more with the unimanual group as a whole, as well as with the entire test, and which determined the 'proper' hand over 80 percent of the time. Five of these tests chosen showed a high consistency (over 90 percent) in Catherine Hull's (7) questionnaire for laterality. The questions thus chosen were modified only to the extent of being made suitable for the present test situation. The sections and numbers of Koch's original test items are indicated below in the brackets following each test's instructions.

After recording the *S*'s age and sex, *E* had the *S* write his name on a piece of paper. The hand that *S* used in this task was noted. *S* was then directed to a table on which the following objects had been laid out: flyswatter, pocket comb, clothes brush, hammer, board with nail in it, whisk broom, dust pan, a piece of gummed paper, and a sheet of typing paper on which the numbers from one to 10 had been written in a random order but with number 10 in the center. *S* approached each object in turn in such a manner that it was equally distant from both hands. Those objects having handles always had the handles pointing directly toward the *S*, rather than to either side. The following instructions were given for each task:

1. "Let's pretend there's a fly here (pointing to spot several feet from *S*). Show me how you'd kill it." [III-6]
2. "Would you show me how you comb your hair?" [IV-2]
3. "Pretend that the front of your clothes are dirty. Show me how you'd brush them off." [IV-4]
4. "Would you hammer this nail into the board for me, please?" [IV-1]
5. "Would you want to brush up the floor for me? I'll hold the dust pan." [IV-25]
6. "Would you stick this piece of scotch-tape over the largest number on this paper?" [III-20]
7. "Show me how you brush your teeth. You may use your fingers." [IV-3] [no materials used]

Once the *S*'s dominant hand had been determined and recorded, he was presented with the mirror drawing apparatus and the star diagram.

The mirror drawing apparatus was similar to those used in the traditional experiments, composed of a vertical mirror, attached to, and held upright by, a flat horizontal base which extended toward the *S* and on which the *S* could write. Attached to an arm extending from the side of the mirror was a screening board which permitted the *S*, sitting directly in front of the apparatus, to view the horizontal base in the mirror, yet be unable to see it directly. This allowed the *S* to trace a design of a star which was placed on the base, screened from his direct view, and visible only in the mirror. To avoid having the *S* hit the screening board with his arm, it was transferred each time to the side of the mirror opposite from the hand being used to draw.

The design used was a six-pointed star, designated by 12 small black dots, 1¼ in. apart. Each of the black dots was in the center of a circle ¼ in. in diameter; a light hyphenated straight line between the dots helped to emphasize the outline of the star. The single points of the star pointed toward the short edge of the paper and the double points of the star were toward the long edge of the paper. When viewed in the mirror, there were four horizontal lines and eight diagonal

lines to be made to complete the star; there were no vertical lines. Two stars were drawn on an 8½ × 11 in. sheet of paper, separated by a straight line, and were designated I and II, respectively.

After *S* was seated comfortably before the apparatus, the star diagram was shown directly to him and the following instructions were given:

"Do you see these dots [pointing]? What do they form? What do they look like? [*S* was given opportunity to answer.] That's right; they form a six-pointed star."

"Now what I want you to do is to draw a line from this black dot [indicating inner dot] to this black dot [indicating] and from this black dot to this one, and so forth, all the way around the star." [*E* moved his finger in the direction indicated by an arrow on the paper beginning at the dot furthest from the short edge of the paper.]

"Now you may find it easy, or you may find it hard. In either case, *once you've started, don't say a word, but work steadily to see how many of the black dots you can touch in the time that I give you.*"

"Now *the lines do not have to be straight.* Just connect each one of these black dots in the proper order. Are there any questions, so far?" [If questions were raised, they were answered.]

The sheet of paper with the two star diagrams was placed in position on the apparatus which was directly in front of *S* and *S* was asked:

"Can you see Star Number 1 in the mirror? Can you see all of it?" [*E* used his hands to indicate which star he referred to; if *S* had difficulty in seeing it, the screening board, which was on the side opposite the drawing hand, was adjusted.]

"Now, you're going to do this with both hands, but since the stand is already set up this way, you can start with your ——— hand first."

"You'll start here [*E* adjusted pencil for *S* on the starting dot and held it there while completing the instructions. The starting dot was the one furthest from the mirror in the star nearest the mirror] and move in the direction of that arrow" [pointing to an arrow drawn between dots one and two to indicate the direction *S* should go].

"Hold the paper down here [indicating location below the dividing line between the two stars] so it won't slip."

"Now I want you to look only in the mirror and when I say 'go' work steadily to see how many of the black dots you can connect in the time that I give you."

"*Keep your pencil on the paper at all times*, and be sure that you *touch each of the little black dots.* Are there any questions?"

"Ready? Shut your eyes. [*E* waited until two seconds before starting time and then said:] Open your eyes. (*E* moved his own hand back and forth between *S* and mirror.) Go." [*E* removed his hand.]

As *S* worked, *E* observed to see that the instructions were obeyed. Many *S*s entered the test situation with confidence, only to become quite upset when they were unable to control the movement of their hand. This resulted in a sudden cessation of work, in an attempt to move the pencil through the air rather than along the paper, or in a misdirection of attention from the problem to the *E*. Also, once *S* was able to move in the correct direction, especially after a struggle, he was often reluctant to slow his movements enough to touch the black dots as directed. If he completely missed the circle around the dot he was immediately told to return and to touch the dot, but if he succeeded at least in touching the circle, he was merely warned at the end of that trial to touch the black dots in the future. At any time during the experiment, *E* repeated any part of the instructions which it seemed necessary to remind the *S* to obey. The words, "Make a mark at all times," were occasionally substituted for the instructions to keep the pencil on the paper. Only in a few cases was there any difficulty after the first trial had been completed with each hand.

The *S* was given one minute in which to draw the star. If he finished sooner, the exact time in seconds was noted. If he did not finish within the minute's time, the point where the pencil mark ended was noted.

After the *S* completed a trial with one hand, the *E* used encouraging words whenever necessary and then reversed the position of the design on the mirror apparatus, reversed the screening board, and told *S* to repeat the test with his other hand. The same procedure was then followed as previously. Only those instructions which had been disobeyed were given again. New diagram sheets were supplied, and the entire procedure was repeated until the *S* had had four one-min. trials with each hand, alternatingly.

In order to control the variables of order of hand performance, the direction moved, and handedness in general, each was varied systematically. For the 40 right-handed *Ss*, 20 used their right hands first and 20 used their left first. Of each of these groups of 20, 10 moved in a counter-clockwise and 10 moved clockwise for every trial. A similar arrangement was made with the 20 left-handed *Ss*, with 10 *Ss* beginning with their left hand, 10 with their right hand, and within each of these groups of 10, five always moved in a clockwise direction and five always moved in a counter-clockwise direction. These variables were indicated on each performance sheet and properly incorporated into the instructions. Variables such as age and sex, which have been shown to affect mirror drawing performance (3), were balanced as well as possible. Little importance needs to be attached to these variables, however, for each *S* represented his own control, comparison being made between his own right and left hand performance.

Scoring was done in the following manner. If the *Ss* failed to finish the star in the minute's time, his score was the number of outer circles reached beyond the starting dot, plus tenths of points for progress beyond the last dot reached. The score was obtained according to the following formula:

$$\text{Score} = \begin{matrix}\text{No. of outer circles reached beyond starting dot.}\end{matrix} + \left(\frac{\begin{matrix}\text{Distance of ending point of pencil from} \\ 10 - \text{next dot, in units of } \tfrac{1}{8}\text{th inch}\end{matrix}}{10} \right).$$

The additional points in the parentheses, if negative, were not subtracted from the number of dots or circles already reached. It thus gave an additional score in tenths of a point from .1 to .9.

However, if the *S* succeeded in completing the star within the minute's time, the number fo seconds that it took him were converted into a score based on his speed, indicating the number of dots he would have reached at the same rate of progress had he used the full minute. The formula would be:

$$\text{Score} = 12 \times \left(\frac{60}{\text{(No. of sec. taken to complete the star)}} \right).$$

Results and Discussion

A glance at the mean scores for each hand of right and left-handed *Ss* in Table I shows a consistant superiority of the non-

TABLE I

The Comparison of Mean Performance Scores between Dominant and Non-Dominant Hands of 40 Right-Handed *Ss*, 20 Left-Handed *Ss*, and all 60 *Ss*, and the Significance of their Differences on Four Trials

Group	Trial	Mean Performance Score		Difference	*t*	*d.f.*	*P*
		Dominant	Non-Dominant				
RH *Ss*	1	8.56	10.29	1.73	2.11	39	<.05
	2	12.35	14.68	2.33	2.84	39	<.01
	3	16.39	18.73	2.34	3.20	39	<.01
	4	19.40	21.14	1.74	2.04	39	<.05
LH *Ss*	1	13.81	16.04	2.23	1.98	19	<.10
	2	18.90	21.34	2.44	1.89	19	<.10
	3	22.70	26.08	3.38	2.56	19	<.02
	4	28.47	29.67	1.20	.98	19	<.40
All *Ss*	1	10.32	12.21	1.89	2.91	59	<.01
	2	14.53	16.90	2.37	3.43	59	<.005
	3	18.49	21.18	2.69	4.13	59	<.001
	4	22.42	23.98	1.56	2.26	59	<.05

dominant hand for every trial. Applying the *t*-test for related means to these results, the differences in favor of non-dominant superiority were significant at better than the .05 level for right-handed *Ss* on all trials, but significant for left-handed *Ss* only on the third trial at better than the .02 level. First and second trials for left-handed *Ss* showed strong tendencies in the same direction, significant at better than the .10 level, but the fourth trial showed no significant difference between hand performances.

Applying χ^2 to these same data, Table II shows that on all trials for both right and left-handed *Ss*, except trial 4 for left-handed *Ss*, the non-dominant hand was superior. The probability of this departure from the null hypothesis of equal distribution of superiority

TABLE II

THE PROBABILITY DETERMINED BY χ^2 OF COMPARED FREQUENCIES OF DOMINANT AND NON-DOMINANT HAND SUPERIORITY FOR 40 RIGHT-HANDED *Ss* 20 LEFT-HANDED *Ss*, AND ALL 60 *Ss* ON FOUR TRIALS *

		Trials			
		1	2	3	4
RH *Ss*	Dominant	12	12	11	11
	Non-dominant	26	26	28	24
	P	0.03	0.03	0.01	0.05
LH *Ss*	Dominant	4	4	4	7
	Non-dominant	15	15	16	9
	P	0.02	0.02	0.01	0.70
All *Ss*	Dominant	16	16	15	18
	Non-dominant	41	41	44	33
	P	0.01	0.01	0.01	0.05

* Sums between dominant and non-dominant hands did not always correspond to the total number of *Ss* used in the experiment, for in most of the trials there were some *Ss* whose performance was equal in the two hands, and so were not counted in this analysis.

in dominant and non-dominant hands was less than 0.05. It is interesting to note that while right-handed *Ss* showed a higher critical ratio between scores than left-handed *Ss* when the *t*-test was used, the probability, as determined by χ^2, of a division in favor of non-dominant hand performance superiority was greater for the left-handed *Ss*.

Tests of independence by χ^2 between the results of right and left-handed *Ss* yielded *P*-values of each trial between 0.40 and 0.50. This indicates that the right- and left-handed groups did not show a significant difference in the relative frequency with which each hand—dominant or non-dominant—is superior. This being the case, the results of all *Ss* were treated together in order to examine performance

differences between dominant and non-dominant hands. Significant differences were yielded by *t*-tests (Table I) in favor of the non-dominant hand rejecting the null hypothesis between 0.01 and 0.03 levels for all four trials. Based on the null hypothesis, χ^2 (Table II) gave *P*-values of less than 0.01 on all trials but the fourth, which was 0.05.

Therefore, we may feel confident that the factors operating which caused the non-dominant hand to be significantly superior to the dominant hand in both right and left-handed individuals in a mirror drawing situation did not occur by chance. Our results become even more significant when we realize that the tests applied in these cases were based on the null hypothesis that there was no difference in superiority between hands. However, the truth of the matter remains that we actually *should* expect a difference in performance between hands in favor of the dominant hand, for by definition and by means of a handedness test, we had chosen *Ss* who showed superior motor ability in their dominant hand. Therefore, theoretically, even if the two hands were equal in ability, it should be looked upon as significant. Actually, however, we have seen that the performance of the non-dominant hand in mirror drawing was significantly better than that of the dominant hand, even when treated according to the hypothesis that there is no difference between hands.

This result is in line with theories of proactive inhibition or negative transfer that have already been discussed. The formation of the new connections is interfered with in its early stages by the appearance of the old and now incorrect responses. This suggests that previous connections built up between the visual and kinaesthetic senses must be torn down to do mirror tracing and replaced by new ones. For the non-dominant hand, these old connections are weakest; this results in superior mirror drawing performance for that hand.

The results in Table I also indicate a consistant superiority in mirror drawing ability for left-handed *Ss* over right-handed *Ss*. One partial explanation for this—in line with the proactive inhibition theory—might be that due to cultural demands upon left-handed *Ss*, they may have a tendency toward ambidexterity. This tendency is not immediately made evident by such tests of handedness as were used in the selection of *Ss* for this experiment. But it may nevertheless be present, and consequently manual habits would not be as strongly developed in both hands of the left-handed *Ss* as in those of the right-handed ones. This, if correct, would make it easier for left-handed *Ss* to break down and replace the old connections with new ones in both hands, thus facilitating the mirror drawing task.

The question next arises: if habit interference is operating, then why did not all of the *Ss* do better with their non-dominant hand? To answer this, we must further examine our data with regards to the other two variables of primary hand performance and direction of movement.

Table III shows the mean differences between right and left-handed *Ss* who began the experiment with their dominant hand and those *Ss* who began the experiment with their non-dominant hand. Positive scores indicate that the difference was in favor of the non-dominant hand.

TABLE III

RESULTS INDICATING EFFECT OF HAND PERFORMANCE ORDER ON MEAN DIFFERENCES BETWEEN HANDS OF RIGHT-HANDED AND LEFT-HANDED *Ss* FOR FOUR TRIALS *

		Trials			
		1	2	3	4
RH *Ss* ($N=20$)	RH first	3.52	4.60	4.06	2.14
	LH first	−.07	.08	.61	1.34
LH *Ss* ($N=10$)	LH first	3.74	4.48	4.81	1.79
	RH first	.71	.40	1.95	.60

* Positive difference indicates non-dominant hand performance superiority.

It is quite clear that when the dominant hand was used first a greater difference was created in favor of the non-dominant hand than when the non-dominant hand was used first. This suggests that learning, or positive transfer, as the case may be, occurs, not only within the same hand, but between hands, from trial to trial. Examining the performance scores in Table I, we can see a steady increase in scores from first hand first trial, to second hand first trial, to first hand second trial, and so forth. Thus, the hand that performed second benefitted from the first hand's performance, so that when the dominant hand went first, the difference between its performance and that of the non-dominant hand which followed was much more in favor of the non-dominant hand than when the reverse performance was used. However, the important point to observe is that except for the first trial of the non-dominant hand of left-handed *Ss*, the overall tendency for the non-dominant hand to be superior was always present. This suggests that the positive transfer, or learning, effect was outweighed by the negative effect of proactive inhibition from the pre-experimental training.

Travis' (11) results, tending already in the direction of non-dominant hand superiority in a mirror tracing task, might have been significantly so, had he systematically varied the beginning hand. As it was, all *Ss* performed with their left hand first, which resulted in a positive transfer effect biasing the results in the direction of the right and dominant hand.

These facts, therefore, strongly suggest that to ignore the effect of positive transfer from the first to the second hand's performance obscures, though never completely eliminates, the action of proactive inhibition on hand superiority in mirror drawing.

Mean differences of right and left-handed *Ss* who moved in a clockwise and in a counter-clockwise direction were compared in a similar manner in Table IV. A positive score indicates a difference in favor of the non-dominant hand.

TABLE IV

RESULTS OF VARYING THE DIRECTION MOVED ON MEAN DIFFERENCES BETWEEN HANDS OF RIGHT-HANDED AND LEFT-HANDED *Ss* FOR FOUR TRIALS *

		Trials			
		1	2	3	4
RH *Ss* ($N=20$)	Counter-clockwise	2.24	2.62	2.82	2.72
	Clockwise	1.21	2.07	1.85	.76
LH *Ss* ($N=10$)	Counter-clockwise	3.84	2.28	4.81	.81
	Clockwise	.61	2.60	1.95	1.58

* A positive difference indicates superiority in the non-dominant hand.

These results suggest that no matter whether the *S* was right or left-handed, there was, in general, a slightly greater difference in favor of the non-dominant hand when the *S* moved in a counter-clockwise direction. This effect was highly irregular for the left-handed *Ss*; this might be accounted for by the smaller number in this group.

One might suppose, however, that the direction of movement would have no effect upon the performance if the *S* is able to complete the star, for between the 12 points, he must make an equal number of hand movements to the right and left, diagonally and horizontally. The differences which did occur, as a result of the direction moved, might be explained by the following hypothesis. Moving counter-clockwise meant beginning by going to the right. The effect of proactive inhibition in this mirror drawing task may have made it easier for the non-dominant hand, as indicated, because

of the older and stronger habits of the dominant hand to move to the right when writing. This would be true for *both* right and left-handed *Ss*. It is during these beginning strokes that many *Ss* gained confidence and rapport as well as insight into the problem of mirror drawing. Thus, ease at this point might psychologically affect the *S*'s entire performance enough to account for a greater superiority in favor of the non-dominant hand when drawing in the counter-clockwise direction. In individual cases, there were reversals which were contrary to the hypotheses established above. It would be impossible to account for all of these reversals. Possibly some of them may be due to an initial motor superiority in holding the pencil in the dominant hand, nervousness on the initial trial, or sudden insight into the task itself.

One factor, important from both a theoretical and a methodological standpoint, should be stressed for consideration in future experiments. In this paper, performance was measured only in terms of time; errors were ignored. Numerous experiments, such as Ewert's (**6**), have demonstrated that the two methods of scoring mirror drawing performance should be considered together. It is quite conceivable, even though quality was not stressed in this experiment, that a person might have attempted to reduce his errors by sacrificing time. W. McTeer (**10**) found that the introduction of electric shock punishment for errors caused a reduction in errors but an increase in time. The quality of the drawings in this experiment showed extreme variability, from almost smooth straight lines between dots to wild uncontrolled lines that moved across the entire star before the succeeding dot was reached. The writer believes that in certain cases where superiority of the non-dominant hand, as measured by time, was not present, it might easily have been present had a measure of errors been used. For example, it is interesting to note that of the 18 *Ss* who spontaneously commented upon the facility of their non-dominant hand in mirror drawing as compared to their dominant hand, one actually did *not* do better with his non-dominant hand by the time measure. This suggests that a design which would measure errors as well as time might yield even clearer results than those obtained in the present experiment. C. E. Lauterbach's (**9**) star design might be modified to fulfill this requirement satisfactorily.

Conclusions

1. In general, the non-dominant hand of both right and left-handed *Ss* was superior in mirror drawing performance to the dominant hand. This is interpreted as an effect of proactive inhibition; the strength of the old connections between hands and eyes interferes with the new connections which must be developed for this task. Assuming that the old connections are stronger for the dominant hand, it is therefore easier to tear down the old and rebuild new connections in the non-dominant hand.

2. There was no significant difference between right and left-handed *Ss* in the relative frequency of the appearance of this phenom-

enon; however, left-handed *Ss* showed a consistant superiority over right-handed *Ss* in ability to do mirror drawing.

3. Learning occurred both between hands and between trials. The hand that performed second received some positive transfer from the hand that performed first. This positive transfer between hands sometimes counteracted the negative transfer effect, but did not completely eliminate it.

4. Moving in a counter-clockwise direction slightly increased the performance superiority of the non-dominant hand over the dominant hand.

(Manuscript received October 29, 1947)

References

1. Betts, E. A. *The prevention and correction of reading difficulties.* Evanston, Ill.: Row, Peterson, and Co., 1936, p. 119.
2. Bray, C. W. Transfer of learning. *J. exp. Psychol.*, 1928, **11**, 443–467.
3. Clinton, R. J. Nature of mirror drawing ability: norms on mirror-drawing for white children by age and sex. *J. educ. Psychol.*, 1930, **21**, 221–228.
4. Cook, T. W. Studies in cross education. I. Mirror tracing the star-shaped maze. *J. exp. Psychol.*, 1933, **16**, 144–160.
5. Cook, T. W. Studies in cross education. II. Further experiments in mirror tracing the star-shaped maze. *J. exp. Psychol.*, 1933, **16**, 679–700.
6. Ewert, P. H. Bilateral transfer in mirror-drawing. *Ped. Sem.*, 1926, **33**, 235–249.
7. Hull, Catherine. A study of laterality test items. *J. exp. Educ.*, 1936, **4**, 287–290.
8. Koch, Helen L. Study of the nature, measurement, and determination of hand preference. *Genet. Psychol. Monogr.*, 1933, **13**, 117–221.
9. Lauterbach, C. E. An improved technique in the mirror tracing experiment. *J. exp. Psychol.*, 1933, **16**, 451–453.
10. McTeer, W. Changes in grip tension following electric shock in mirror tracing. *J. exp. Psychol.*, 1933, **16**, 735–742.
11. Travis, L. E. A comparative study of the performance of stutterers and normal speakers in mirror tracing. *Psychol. Monogr.*, 1928, **39**, No. 2, Whole No. 178, 45–50.

QUANTITATIVE RELATIONS AMONG VERNIER, REAL DEPTH, AND STEREOSCOPIC DEPTH ACUITIES[1]

BY RICHARD N. BERRY

Indiana University

In 1838 Wheatstone wrote, ". . . the mind perceives an object of three dimensions by means of two dissimilar pictures projected by it on the two retinae . . ." (11, p. 373). Although some of the terminology might be challenged, this fundamental tenet has never been seriously questioned. However, it has never been clearly demonstrated that, under comparable experimental conditions, a stereoscopic fusion of the two dissimilar 'pictures' (Fig. 1A) yields the same threshold for depth discrimination as is yielded by a real displacement in depth (Fig. 1B). The test objects used in the present experiment were in the form of black vertical rods appearing one above the other as U and L in Fig 1D. In the stereoscopic situation depicted in Fig. 1A the stereoscopically fused image of the upper rods (U) appeared to be nearer to the observer than the similarly fused image of the lower rods (L), provided that angles 1 and 2 were sufficiently large. In Fig. 1B, rod U was actually nearer to the observer than rod L and so appeared, provided that angles 3 and 4 were sufficiently large. If Wheatstone's statement, quoted above, is to be considered valid, the two methods of presentation, stereoscopic depth and real depth, should have yielded the same threshold values. The present experiment indicates that the two methods of presentation did, in fact, yield thresholds which were almost identical with one another.

A second problem, partially a restatement of the one above, has arisen in regard to liminal depth displacement in its relation to liminal vernier or lateral displacement. A vernier displacement is here defined as a misalignment of objects or edges in the frontal plane (Fig. 1C). It might be expected that the liminal amounts of these vernier displacements for each separate eye would bear a certain relation to the liminal values of those vernier displacements which are binocularly fused to give the impression of depth. Two alternative hypotheses concerning this relation have been commonly made. These

[1] This paper is based on a thesis submitted to the Graduate School of Brown University in partial fulfillment of the requirements of the degree of Doctor of Philosphy. The writer wishes to express his appreciation to Dr. Lorrin Riggs for guidance and material assistance in the conduct of the present experiment. To Dr. Richard L. Solomon and Dr. Carl P. Duncan, as well as to Dr. Riggs, are extended thanks for expending many unremunerated hours as experimental *Ss*.

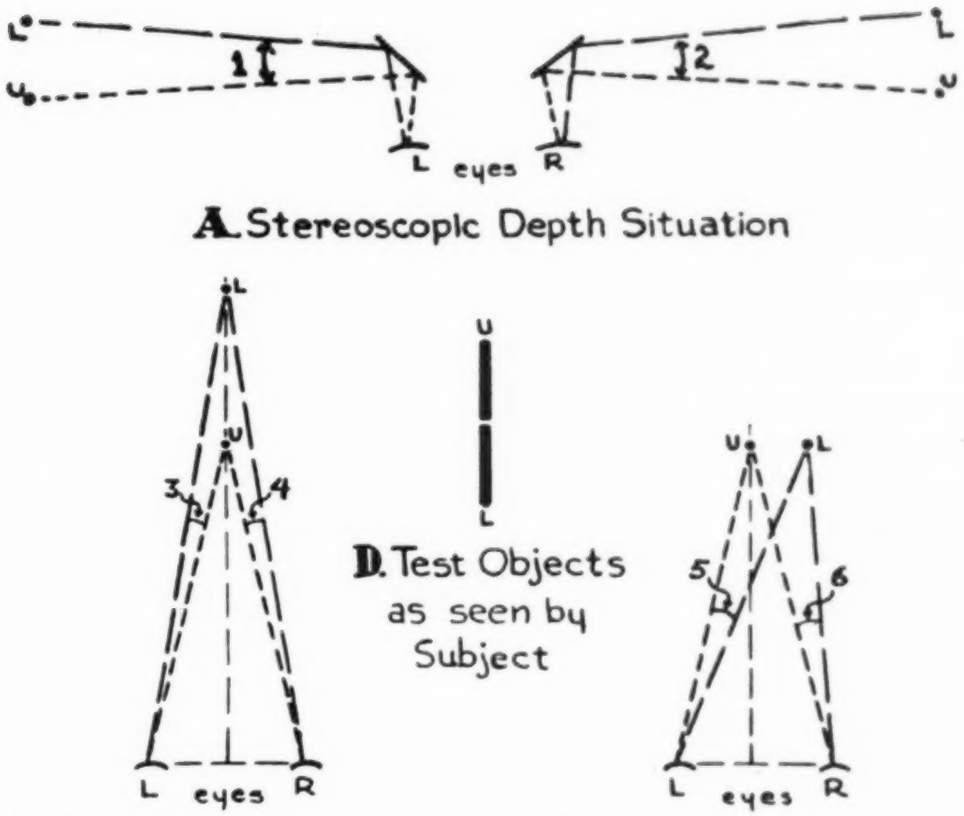


FIG. 1. Schematic representation of the three visual situations employed in the present experiment. U and L represent the upper (stationary) and lower (movable) rods, respectively. Angles 1 to 6 represent threshold displacements of the movable rod in terms of visual angle.

may be expressed by reference to the angles labeled in Fig. 1. All angles are considered as liminal [2] and equal for each eye. One hypothesis then is that:

$$\angle 1 = \angle 2 = \angle 3 = \angle 4 = \angle 5 = \angle 6, \qquad (1)$$

while the other hypothesis is that:

$$\angle 1 + \angle 2 = \angle 3 + \angle 4 = \angle 5 = \angle 6. \qquad (2)$$

The results of the present experiment do not appear consistent with either of the above hypotheses. They show instead an unexpected relation:

$$2(\angle 1 + \angle 2) = 2(\angle 3 + \angle 4) = \angle 5 = \angle 6, \qquad (3)$$

or,

$$\angle 1 = \angle 2 = \angle 3 = \angle 4 = \angle 5/4 = \angle 6/4.$$

This formulation does not apply to some of the rather special

[2] The thresholds herein discussed are in the usual units of visual angle. This angle corresponds to the component in the frontal visual plane of any displacement in relation to the distance of the test object from the eyes.

conditions that were employed in the present experiment, but is offered as a broad generalization of our experimental findings.

The vernier and real depth portions of the present experiment are based on a suggestion made by Walls (10). Walls realized that vernier thresholds and depth thresholds had never been determined under comparable conditions, thus making any comparisons between the two very difficult. He therefore proposed the construction of a test object which could be used for the determination of both thresholds. The proposed 'stereo-vernier' test object involved two rods placed vertically one above the other as the upper and lower objects, U and L, in Fig. 1D. One of these rods was to be the variable and capable of excursion in either the median ('stereo') plane or the frontal ('vernier') plane. The other rod was to act as the standard. Since, in our experiment, it was not found mechanically feasible to incorporate both types of motion into one piece of apparatus, one mechanism was devised for driving the variable rod in the median plane, while another mechanism was designed for moving this rod in the frontal plane. The net result was the creation of a situation such as Walls had proposed.

The stereoscopic portion of this experiment was added as a needed empirical check upon the geometrical equivalence of real and stereoscopic depth. The outcome of this portion of the experiment is significant in showing, in the absence of other cues, the relative effectiveness of stereoscopic disparity in the judgments of real and stereoscopic depth.

Historical

There have been many determinations of both vernier and real depth thresholds. The quantitative studies on stereoscopic depth are far more difficult to perform and appear to be much less numerous.

Apparently the first measurement of vernier thresholds was performed by Wuelfing (13). He noted that in previous experiments the test objects had always been placed in a lateral relation to one another, as parallel lines, rather than one above the other. He attributed the large thresholds which had been obtained by the use of such parallel lines to the functioning of irradiation on the retina. Wuelfing reasoned that if the vertical lines were not placed next to one another, but rather vertically one above the other, the irradiation would be precluded. When such conditions were actually imposed, the vernier threshold was found to be about 12″ of visual angle. Stratton (9) found a vernier threshold of 7″, and Baker and Bryan (3) obtained values as low as 3.7″. Averill and Weymouth (2) more recently reported values of slightly less than 2″. French (5) seems to be the only investigator to report values as low as 0.45″, which, in the light of results obtained by others, would appear to be unaccountably low.

Depth thresholds have been determined many times since Helmholtz's (6) classical three-needle experiment, from which he reported a threshold

of 60″. Bourdon (4) repeated the three-needle experiment many years after it was originated and obtained thresholds as low as 5″. Following Bourdon, several investigators, among them Howard (7), Andersen and Weymouth (1), Langlands (8), and Woodburne (12), have reported values of about 2″. This correspondence of results reported by the recent investigators is somewhat surprising when it is considered that neither apparatus nor procedure was duplicated from one experiment to another.

In a study of stereoscopic depth thresholds, Langlands (8) utilized some plates designed by Pulfrich and found that the best of ten observers achieved a threshold of 25″; one of the same 10 *S*s obtained a threshold of 1.8″ on the Howard-Dolman apparatus.

Comparisons among thresholds as they are obtained in different experiments by different investigators should be made only with extreme caution. However, it would seem safe to assert that, in general, vernier thresholds have been found to be at least as large as real depth thresholds and usually somewhat larger.

Apparatus

The apparatus of the present experiment was designed for the purpose of presenting vernier, real depth, and stereoscopic depth tasks under as nearly identical conditions as possible. Let us first consider the test objects used in all the tasks. These were vertical steel rods 2.4 mm. in diameter, covered evenly with flat black paint, and placed vertically one above the other. They appeared as black line silhouettes, each rod having an apparent thickness of 107″ of arc. The lower rod subtended 40′ 26″ in the vertical plane, and the upper rod subtended somewhat more or less than this depending on the vertical separation between the upper and lower rods. The top rod served as the fixed standard, and the lower rod served as the variable which was judged in its relation to the fixed rod. The appearance of the rods in the visual field is shown in Fig. 1D. In all cases the fixed rod was fastened securely to a framework which in turn was fastened to a base plate on which the mechanism carrying the variable rod was also mounted. The rods were viewed through apertures in such a manner that the only contours in the visual field were those created by the rods. The rest of the field was uniformly painted with white casein paint and illuminated by batteries of frosted incandescent lamps. None of these lamps was visible to the *S*. The *S* viewed the test objects, in all situations, at a visual distance of 4622 mm.

Let us examine Fig. 2. The test object was seen through two round apertures (A_r and A_l) in the front of a barrel (B) into the other end of which the subject looked. This barrel was uniformly illuminated by a single frosted lamp. One of the round apertures in the front of the barrel (A_r) permitted the right eye to see the test object at point X, while the second aperture (A_l) performed the same service for the left eye. These apertures were 331 mm. in front of the eyes. Their diameters subtended a visual angle of slightly less then 3°. The second aperture, through which each eye viewed the test object, was a single rectangular aperture in surface S_1 450 mm. in front of the test object. This aperture was 1° 23′ in height and 20′ in width. The test rods appeared in the middle vertical plane of this aperture. They were viewed against an illuminated background provided by surface S_2. Great care was exercised in adjusting the barrel and all the other surfaces to equal levels of brightness. In this way the contours of the various apertures were 'washed out' to all but a negligible extent, and the test rods appeared sharp and clear against a uniform white background which involved almost all of the visual field.

A. The vernier situation.—The general vernier visual situation is diagrammed in Fig. 2. This diagram is not drawn to scale but is intended to represent only functional relations. A_s and C_s represent large screens which ran from the floor to the ceiling and prevented stray light from impinging on the *S*'s eyes. The point marked X at the intersection of the dashed lines represents the point where the fixed rod of the test object was placed. The variable rod was carried by a slider which was driven by a micrometer in such a manner that it was movable to the right or left in the same frontal plane as the fixed rod. In front of X there was a mirror shutter M which,

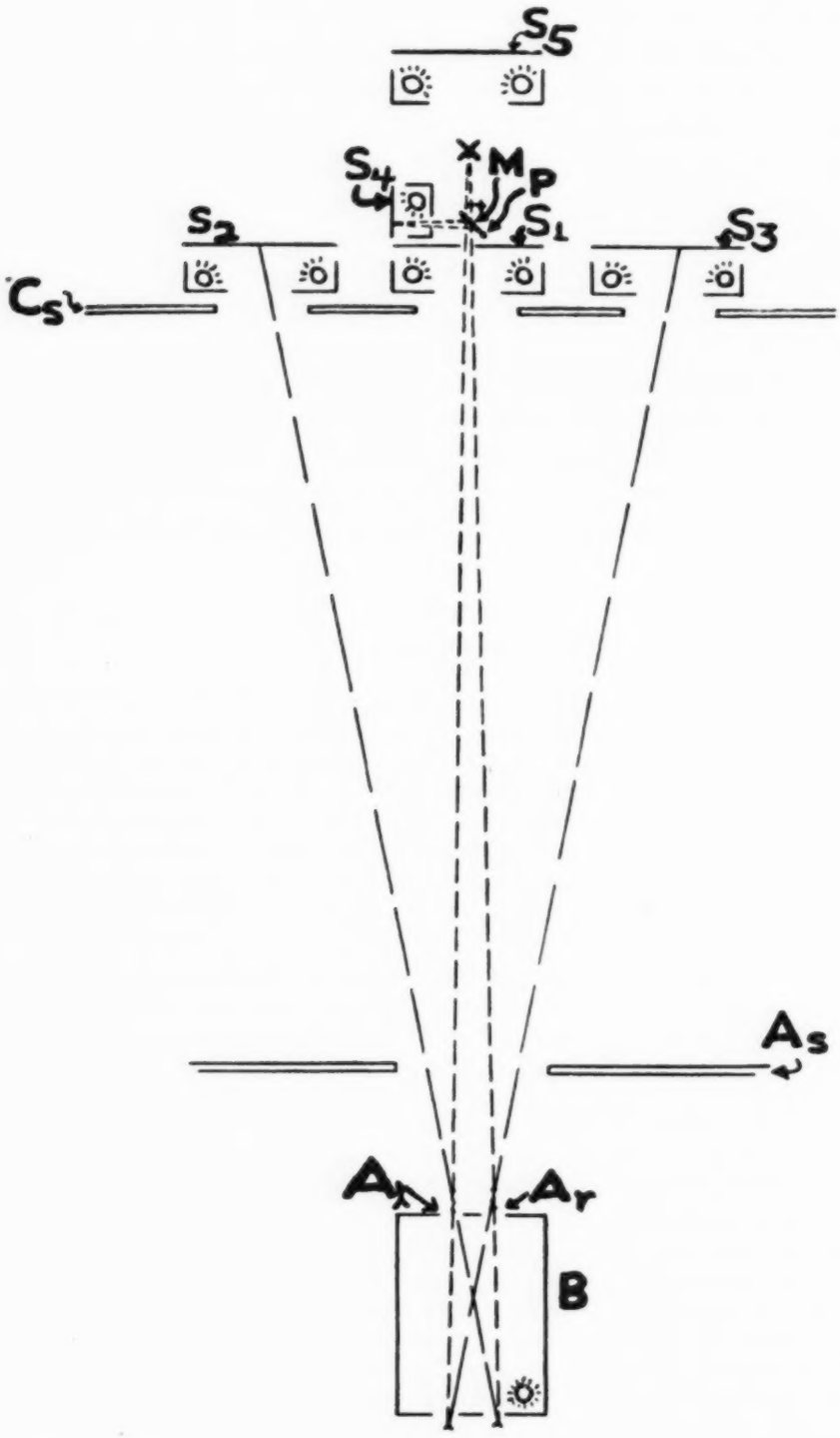


FIG. 2. Diagrammatic representation of the vernier visual situation and the real depth visual situation

when in the position shown, permitted the *S* to view two small fixation crosses printed one above the other on surface S_4. The region between these crosses served as a fixation point which appeared to be replaced by the test rods when the mirror was swung out of the visual field by the activation of an attached solenoid. Complications due to after-images were minimized by the fact that the *S* fixated on a blank region between the crosses. Thus, the *S* was at all times presented with a fixation point at a visual distance of 4622 mm.

The illuminated surfaces S_2 and S_3 furnished illumination for those parts of the retinae which were not illuminated by other sources. The movement of the variable rod was checked by a photographic method and was found to have a negligible error[3].

B. The real depth situation.—The real depth apparatus was placed so that the fixed rod of the test object came at the point X in Fig. 2, and the general visual situation was the same as that for the vernier test object. The real depth test object provided a movement of the variable rod toward or away from the *S* in the median plane in which the fixed rod was also placed. Rods of the same dimensions were employed as had been previously incorporated in the vernier test object. The movement of the variable rod in the median plane was obtained through the operation of a large screw mechanism. The settings were made by means of a pointer running along a scale and could be made with a very small error. Lateral movement of the variable test rod was precluded by an arrangement involving two hardened steel straight edges which served as a track to control the lateral movement of the rod. These straight edges had a tolerance of 0.0001 in. in six in. (Brown and Sharpe Mfg. Co. specifications.) The extent of lateral movement was checked by photography and found to be unimportant.[4] In the real depth situation the *S* was supplied with a biting board to preclude head movements and thus preclude movement parallax.

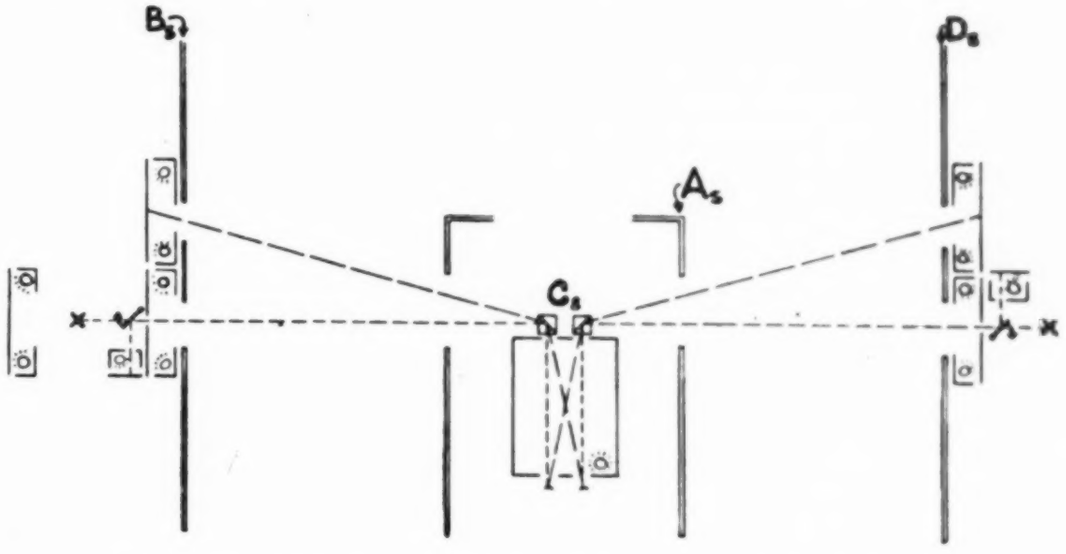


FIG. 3. Diagrammatic representation of the stereoscopic depth situation

C. The stereoscopic depth situation.—The stereoscopic depth situation is diagrammed in Fig. 3. The fixed rods of two vernier test objects were placed at the points of the right and left X's, one fixed rod at each point. The variable rods moved perpendicularly to the visual axes represented by the shorter dashed lines. The micrometers, each attached to one of the variable rods, were connected by a pulley in such a manner that a movement of one rod caused an equal and opposite movement of the second rod. The rods were set by means of a photographic calibra-

[3] A photographic technique was used to test the alignment of the two test rods and the linearity of excursion of the variable rod. A vertical straight-edge was placed close to the rods, and a point source of light was placed exactly in the median plane at a point near the observer's position. The shadows of the rods and the straight-edge, cast by the point source, were allowed to fall on a strip of Kodak micro-film having a resolution of 160 lines per mm. The film negatives were enlarged 60 times by projection on a screen for purposes of measurement.

In the case of the vernier movements (including the two vernier devices used in the stereoscopic situation of Fig. 3) it was found that any desired setting could be made with a maximum error of 0.01 mm., or one-fourth of the smallest constant stimulus interval that was employed. In the case of the real depth test object the lateral movement of the variable rod was of the same magnitude. Since this amount of lateral deviation was well below the threshold of lateral displacement, it can safely be stated that the *S*s were not using such a lateral deviation as a cue in the real depth situation.

[4] See footnote 3.

tion,[5] so that when one had a zero lateral deviation from its corresponding fixed rod, the second also had a zero deviation from its fixed rod. Thus, at any setting an equal and opposite displacement was presented to each eye.

The test objects were viewed through half reflecting right angle prisms placed directly in front of the circular apertures of the barrel at C_e in Fig. 3. These prisms were precision-made optical surfaces designed for use with a 24-power telescopic instrument.

The shutters in the stereoscopic situation were operated simultaneously by the simple expedient of wiring their actuating solenoids in parallel so that when one was activated, the other one was also activated. The visual distance of the test objects from the eyes (4622 mm.) was the same as had previously been employed in the vernier and real depth situations.

D. Monocular situation.—Monocular thresholds for each eye were determined for each of the three principal visual situations that have been described above. In all cases the *S* was instructed to respond to the vernier displacement rather than to any depth displacement that might exist; this was true even when the real depth test object was employed. As for the eye which was not being tested, it was thought best to avoid the complications inherent in occluding that eye. Hence the unused eye remained open and was allowed to view, through a reflecting prism, the fixation crosses in either the right or left field of the stereoscopic situation, depending on which eye was being tested. The eye being tested viewed the corresponding fixation crosses for the particular test object being used. The two sets of crosses were fused without difficulty. The solenoids which actuated the mirror shutters were wired in parallel so that when the test object was substituted for the crosses for the eye being tested, the crosses presented to the other eye simultaneously gave place to a homogeneous white illuminated field from which the test object had been removed. Thus, the unused eye was without a fixation point only during the brief interval when the test object was presented.

Procedure

A. General.—Three male *Ss* with a minimum acuity rating of 20/20 for each eye were employed in the present experiment. The only serious visual defect was an extreme exophoria in Subject L. A. R. All of the *Ss* had previously acted as observers in studies on stereoscopic acuity. The *Ss* were two male graduate students (R. L. S. and C. P. D.) and one faculty member (L. A. R.), ranging in age from 25 to 34 years.

The *Ss* were run on a routine schedule of 100 responses a day. Daily sessions were about 40 min. in length and contained five rest periods. Two of these daily sessions yielded the data for the computation of a single threshold.

The psychophysical method of constant stimuli was used in all determinations of thresholds. Five positions of the variable rod were used in the determination of any given threshold. The variable rod was adjusted to each of the five positions by the use of the apparatus described above. The variable rod was always stationary for any given observation, since lateral or depth 'movements' described above were made by the *E* in adjusting the rods between observations. The sequence in which these positions were presented was determined by reference to a coded order which was prepared at the beginning of the experiment. This meant that the same *order* of stimulus changes was used in all daily sessions. The daily series of 100 presentations yielded 20 responses to each position of the variable rod. Thus for the two daily sessions required to determine a threshold there were 40 responses to each of these positions.

On the first day of a threshold determination, 10 to 15 preliminary responses were obtained at various scale settings in order to determine the five positions of the variable rod that were to be used that day and the next day. Then the series was presented. Each presentation was four sec. long and followed by a 10-sec. interval in which the *S* fixated the area between the fixation crosses in the visual field. Each stimulus presentation was preceded by a verbal 'Ready.' A 'forced choice' procedure was employed wherein the *S* was allowed to make at each exposure only one of two responses. In the real depth or stereoscopic depth situations, the response was either 'Back' or 'Front.' In the vernier situation it was either 'Right' or 'Left.'

The independent variable that was studied in each of the three main experimental conditions was the vertical separation between the upper and lower rods. In the case of the vernier threshold determinations, these separations corresponded, in seconds of arc, to 3.6″, 20.1″, 44.6″, 133.7″, 312″, and 891″ (0.08, 0.45, 1.0, 3.0, 7.0, and 20.0 mm.). In the determination of the real depth and stereoscopic depth thresholds the 3.6″ (0.08 mm.) separation was not employed.

[5] See footnote 3.

The thresholds were first obtained for all vertical separations of the test rods on the vernier test object. Next, the thresholds for all vert al separations of the test rods in the real depth situation were determined, and lastly all stereoscopic thresholds were measured.

The effects of practice were not systematically checked. However, the two checks that were made, on the real depth and stereoscopic depth situations, indicated that practice played a negligible role.

B. Computational.—Thresholds were computed from the raw data by a graphic method. In this method only one of the two categories of response was used in the computations. In the vernier situation only the 'Right' responses were considered, while in the depth situation only the 'Front' responses were used in the computations.[6]

On the basis of sample plots, it was believed safe to assume that percentages of 'Right' or 'Front' responses, when plotted on a probability grid as a function of the displacement of the variable rod, yielded a straight line. On the basis of this assumption the percentages of 'Right' or 'Front' responses obtained for each displacement of the variable rod was determined and the scale readings corresponding to 25 percent and to 75 percent were read from the graph. The difference between these two values was divided by two, thus yielding the conventional threshold, in raw scale units, for the displacement of the variable rod. This amount was then converted, in each case, into units (seconds of arc) corresponding to the threshold visual angles shown in Fig. 1.

Results

The thresholds for individual *S*s for each principal experimental condition may be found in Table I and Fig. 4. It is apparent from these data that the stereoscopic and real depth thresholds are quite

TABLE I

Vernier, Real Depth, and Stereoscopic Depth Thresholds of Each Subject for Each Vertical Separation of the Test Rods *

Visual Situation	Subject	Vertical Separation					
		0.08	0.45	1	3	7	20
Vernier	R. L. S.	1.67	1.32	1.23	3.44	3.22	5.93
	C. P. D.	2.31	1.62	2.18	4.50	4.15	6.73
	L. A. R.	2.70	1.36	1.65	3.12	3.37	6.55
Real depth	R. L. S.		1.50	1.55	1.69	1.13	2.64
	C. P. D.		2.88	2.94	2.35	1.52	2.67
	L. A. R.		2.26	2.94	2.56	2.94	3.39
Stereoscopic depth	R. L. S.		1.25	1.82	2.04	1.64	3.82
	C. P. D.		4.56	3.12	2.53	2.14	3.12
	L. A. R.		2.83	3.14	2.82	2.49	4.00

* All thresholds are in seconds of arc. The vernier thresholds refer to the size of angle 5 in Fig. 1; the real depth thresholds refer to the sums of angles 3 and 4; and the stereoscopic depth thresholds refer to the sums of angles 1 and 2.

similar for each *S* and are similarly affected by changes in the vertical separation between the test rods. The vernier thresholds, on the other hand, are quite differently affected by the factor of separation.

[6] The use of the other categories in each case would have yielded the same function, since the percentage of responses of one category is always equal to 100 minus the percentage of responses of the other.

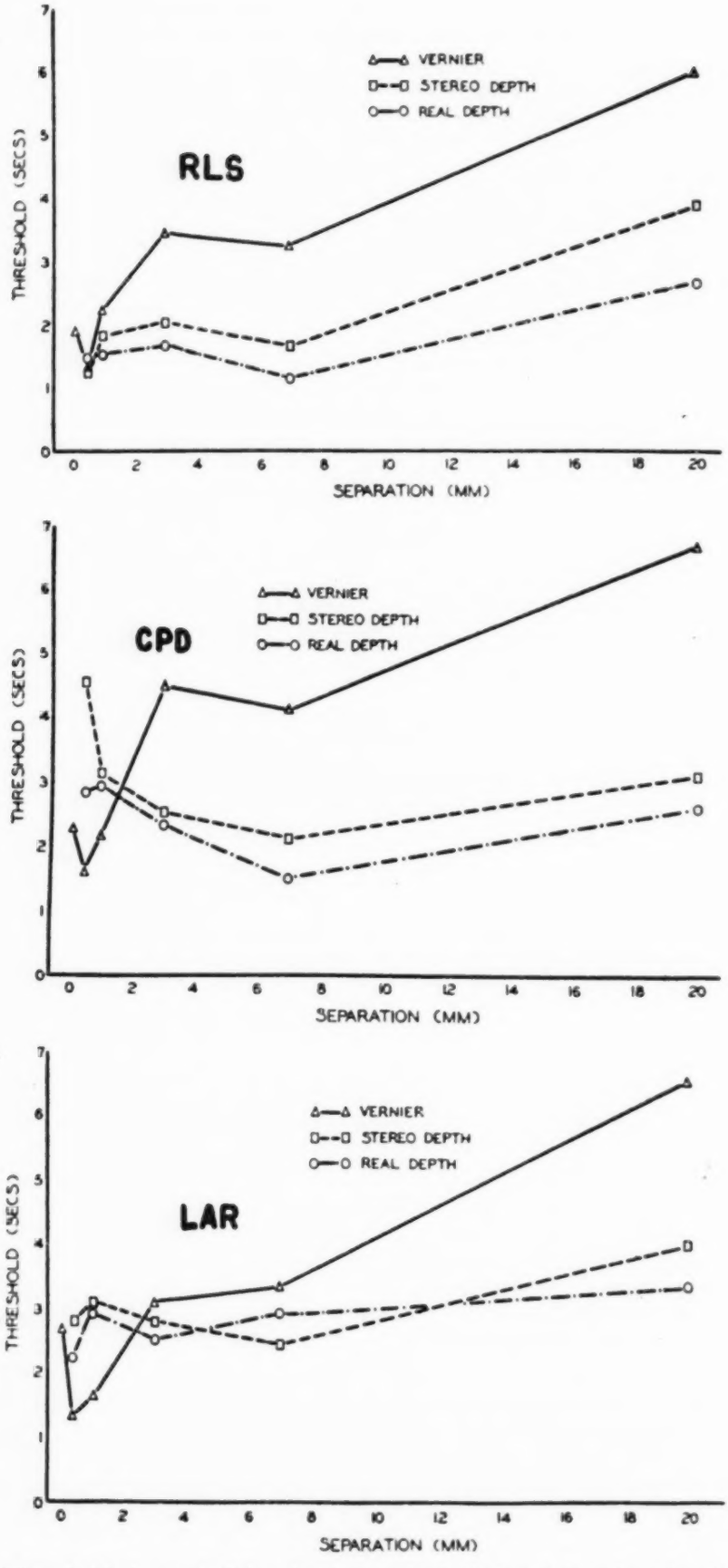


FIG. 4. Thresholds for each *S* for the three principal visual situations for all vertical separations of the test rods

The optimal separation for vernier discrimination appears to be less than one mm. or 44.6 seconds of arc. For the stereoscopic and real depth discriminations, on the other hand, the optimal separation is in the neighborhood of seven mm. or 312 seconds. Furthermore, the vernier threshold is greatly influenced by the factor of separation; and the characteristic shape of this function is consistent from one *S* to another. For all separations of three mm. or more, the vernier thresholds are progressively higher than the corresponding real and stereoscopic depth thresholds. It seems safe to predict that the thresholds of all types would continue to rise indefinitely if the separations were increased beyond those that were employed in these experiments.

All the results that have so far been mentioned were obtained under conditions where the test object was viewed with both eyes, and judgments were made of lateral displacement (vernier situation) or apparent distance (real depth and stereoscopic depth situations). Thresholds were also obtained with monocular observation, in which case the appearance of depth was lacking and all judgments were of a verneroid displacement. The response was either 'Right' or' Left.' Thresholds for each eye were obtained for all *Ss* in the three test situations that were employed. In all cases the eye not viewing the test object was illuminated as described in the section on procedure. In these monocular determinations only the three-mm. vertical separation of the test rods was used.

TABLE II

MONOCULAR AND BINOCULAR THRESHOLDS AT THE THREE MM. VERTICAL SEPARATION

		Visual Situation	Subjects		
			R. L. S.	C. P. D.	L. A. R.
A. Monocular thresholds*	Right eye	Vernier	3.03	4.37	3.18
		Real depth	2.44	3.40	3.29
		Stereoscopic depth	2.16	5.36	3.22
	Left eye	Vernier	3.21	3.53	3.52
		Real depth	3.44	3.23	3.01
		Stereoscopic depth	4.00	3.12	4.54
B. Binocular thresholds**		Vernier	3.44	4.50	3.12
		Real depth	1.69	2.35	2.56
		Stereoscopic depth	2.04	2.53	2.82

* All responses were either 'Right' or 'Left.'
** Taken from Table I.

All thresholds for monocular determinations are presented in Table II. From this table it may readily be seen that the monocular threshold varied slightly from one determination to another. No

consistent tendency, however, is evident for any one procedure to yield thresholds lower than those obtained by any other procedure. Therefore, it can be said that all methods of measuring the threshold of vernier acuity yielded essentially the same value for any given subject. In Table II may also be found the thresholds obtained under all binocular conditions for the three-mm. separation. A comparison between the monocular and binocular *vernier* thresholds indicates that the two values are approximately the same for any given *S*.

Reliability of any given threshold is not accurately known, since no threshold for a given experimental condition was determined more than twice. However, with a three-mm. vertical separation of the test rods, thresholds were determined twice for the real depth situation and twice for the stereoscopic depth situation. In both cases, the second determination has been the one used for further computations. A comparison between the first and second determinations may be made from an examination of Table III. It is evident from such a

TABLE III

REAL DEPTH AND STEREOSCOPIC DEPTH THRESHOLDS FOR EACH OF TWO DETERMINATIONS

		R.L.S.	C.P.D.	L.A.R.
Real depth thresholds	First determination	1.37	1.37	*
	Second determination	1.32	2.35	2.15
Stereoscopic depth thresholds	First determination	1.67	2.23	3.91
	Second determination	1.84	2.01	3.73

* This threshold could not be determined because the five positions of the variable rod were not sufficiently far apart.

comparison that the differences between the determinations are not great, nor is one determination consistently larger or smaller than the other.

Constant errors were computed only for the real depth and the stereoscopic depth thresholds. The constant errors showed no consistent trends and varied in size from 0 to 14.5″, being slightly larger for the stereoscopic depth situation.

DISCUSSION

The most clear-cut result of the present experiment is that the real depth and the stereoscopic depth situations yield essentially the same thresholds. The evidence indicates that the slight, but consistent, difference between the two sets of thresholds is of a magnitude to be expected from the operation of a slight aberration or misalignment of the prisms used in the stereoscopic situation.

It has been stated that the thresholds for stereoscopic depth and for real depth were almost identical for all vertical separations of the test rods. However, the relation between the vernier thresholds and the depth thresholds differed as the vertical separations were varied. This is interesting when it is considered that the depth situations apparently involved the binocular fusion of two different vernier displacements.

Let us consider Fig. 1 for a moment. It has been established, in the present experiment, that for the vernier situation the threshold angle 5 or angle 6 remains the same whether one or two eyes are employed. Thus, there is not a binocular summative effect when the two images on the retinae are identical. However, when these images are not identical, as in the real depth or stereoscopic depth situations, the binocular fusion does result in a summation effect. With the widest vertical separations this summation results in a depth threshold angle 1 (or 2, or 3, or 4) which is one-fourth that of the vernier threshold angle 5 (or 6). With the narrower vertical separations, angle 1, in some cases, is equal to angle 5.

We may now ask a question concerning these relations. Why does the binocular fusion of two identical vernier displacements, as in the binocular vernier situation, yield a larger threshold angle 5 than the threshold angle 1 obtained by the binocular fusion of two disparate vernier displacements? There is no ready answer to this question to be found from the results of the present experiment or other previous experiments. There is only an appeal to some kind of cortical interaction. This interaction would appear to achieve its greatest importance in the region where the vernier discrimination has become rather poor, i.e., at the widest vertical separations of the test rods.

Depth discrimination tends to break down as the vertical separations decrease in size from the optimal separation. This breakdown with decreasing separations tends to occur at somewhat larger separations than does the deterioration of the vernier performance. This is probably attributable to the importance of ocular muscle balance in the depth situations. A muscular imbalance would certainly increase the difficulty of achieving a clear binocular image of a very small separation. The vernier discrimination would not be affected by an imbalance, because the vernier performance is essentially monocular and does not depend on the cooperation of the two eyes. Vernier discrimination would be expected to suffer, however, when the separation becomes so small as to disappear altogether; and this situation is present with the vertical separation of 0.08 mm. With this separation all three *Ss* experienced some difficulty in making vernier judgments and attained thresholds which were somewhat

higher than those obtained at the clearly visible separation of 0.45 mm. between the test rods.

Summary

Three visual discrimination situations were presented, to each of three *S*s. The situations were: vernier, real depth, and stereoscopic depth. In each situation the experimental variable was the vertical separation of the test rods, placed vertically one above the other. The brightness and color of the general visual field and the visual distance, size, and shape of the test objects were identical for all situations. Six different vertical separations were used: 3.6″, 20.1″, 44.6″, 133.7″, 312.0″, and 891″ of visual arc. The visual distance from the eye to the test objects was 4622 mm. in all cases. Monocular vernier judgments were made in all three of the principal visual situations at the 133.7″ separation of the test rods. The stimulus duration for each judgment was four sec. The psychophysical method of constant stimuli with two categories of response was employed.

1. Thresholds obtained in the vernier situation decreased from 6″–7″ at the widest vertical separations to slightly over 1″ at the next to the narrowest separation. The narrowest separation yielded thresholds from 2″–3″

2. Monocular vernier thresholds were, in general, equal to binocular vernier thresholds.

3. The real depth thresholds remained generally within a range from 2″–4″ for all vertical separations with minimum values most frequently at the intermediate separations.

4. The stereoscopic depth thresholds paralleled the real depth thresholds for all separations, but were very slightly (about 0.2″–0.6″) higher. It is assumed that the slightly higher thresholds for stereoscopic depth are accounted for by the necessity for using prisms in this situation, while no optics were used in the remainder of the observations. Hence, the results are taken as evidence that real depth thresholds and stereoscopic thresholds are essentially identical.

5. Real depth and stereoscopic depth thresholds are, for the most part, too low to be accounted for by a simple summation of the two separate monocular, vernier components.

(Manuscript received December 18, 1947)

References

1. Andersen, E. E., & Weymouth, F. W. Visual perception and the retinal mosaic: I. Retinal mean local sign: an explanation of the fineness of binocular perception of distance. *Amer. J. Physiol.*, 1923, **64**, 561–594.

2. Averill, H. L., & Weymouth, F. W. Visual perception and the retinal mosiac: II. The influence of eye movements on the displacement threshold. *J. comp. Psychol.*, 1925, **5**, 147–176.

3. Baker, T. Y., & Bryan, G. B. Errors of observation. In: *Proc. opt. Conv., London.* Hodden and Stoughton, 1912, **2**, 253–265.
4. Bourdon, B. L'acuité stéréoscopique. *Revue philosophique*, 1900, **49**, 74–78.
5. French, J. N. The unaided eye: Part III. *Trans. opt. Soc., Lond.*, 1919–1920, **21**, 127–156.
6. Helmholtz, H. *Treatise on physiological optics*: III. *The perceptions of vision.* (Trans. by J. P. C. Southall.) Ithaca, N. Y.: Opt. Soc. Amer., 1925.
7. Howard, H. J. A test for the judgment of distance. *Amer. J. Ophthal.*, 1919, **2**, 656–675.
8. Langlands, N. M. S. Experiments on binocular vision. *Med. Res. Council (Brit.), Special Report Ser.* No. 133. London: H. M. Stationery Office, 1929.
9. Stratton, G. M. A new determination of the minimum visible and its bearing on localization and binocular depth. *Psychol. Rev.*, 1900, **7**, 429–435.
10. Walls, G. L. Factors in human visual resolution. *J. opt. Soc. Amer.*, 1943, **33**, 487–505.
11. Wheatstone, C. Contributions to the physiology of vision: I. On some remarkable and hitherto unobserved, phenomena of binocular vision. *Trans. Philos. Roy. Soc., Lond.*, 1838, **128**, 371–394.
12. Woodburne, L. S. The effect of a constant visual angle upon the binocular discrimination of depth differences. *Amer. J. Psychol.*, 1934, **46**, 273–286.
13. Wuelfing, E. A. Über den kleinsten Gesichtswinkel. *Z. Biol.*, 1892, **29**, 199–202.

THE EFFECT OF DIFFERENTIAL EXTINCTION ON SPONTANEOUS RECOVERY [1]

BY ALVIN M. LIBERMAN

Wesleyan University

Introduction

It has been suggested (4) that the broad transfer of experimental extinction may be a significant factor in the occurrence of spontaneous recovery. This suggestion was made in an attempt to identify spontaneous recovery as the rapid forgetting of experimental extinction, and was based upon considerations which may be summarized as follows. There is evidence to justify the tentative conclusion that extinction (negative learning) [2] often transfers broadly—more broadly, in any event, than the corresponding positive learning—and secondarily reduces the strength of a large and heterogeneous group of responses. The probability is high, therefore, that the extinction or negative learning will be interfered with in a comparatively short time, since a few of the numerous secondarily extinguished responses will almost certainly be strengthened during any 'rest' interval, and the transfer of this strengthening back to the originally extinguished response will cause the extinction to be forgotten. Positive learning is better remembered, because the relatively more restricted transfer of the original learning, in combination with other factors, sets up a smaller number of everyday activities which have a capacity for interfering with the retention of the originally learned response.

It will be the purpose of this paper to suggest that the broader transfer of extinction, and the spontaneous recovery which is associated with it, may result from certain procedures which characterize negative learning and normally distinguish it from positive learning. The attempt will be made to test this suggestion by determining

[1] The Institute of Human Relations at Yale University generously provided equipment, animals, and research assistance for the larger part of this study.

The writer is very much indebted to Mrs. Claudine Wirths and Mr. Russell Clark for thier suggestions and competent assistance in connection with this research.

[2] It will be assumed that experimental extinction is a learned modification of behavior in the direction of longer latency, smaller amplitude, or less frequent occurrence of the response. For convenience, the type of learning represented by experimental extinction will frequently be referred to in this paper as 'negative' learning. The qualification 'positive' will be applied to those types of learning, e.g., conditioning, instrumental learning, in which the learning is measured in terms of shorter latency, greater amplitude, or more frequent occurrence of the response.

whether the amount of spontaneous recovery is reduced when experimental extinction is carried out according to procedures which are calculated, in terms of an analogy with positive learning, to restrict the transfer of the extinctive tendency.

A comparison of the procedures under which positive learning and extinction normally occur reveals a difference in the degree to which the animal is made to differentiate the stimulus and response aspects of the situation. It is the very essence of positive learning that the animal is required to learn which response to make to which set of stimuli; the experimenter's demands are typically not satisfied if the animal learns merely to respond. In conditioning, for example, the animal is forced to isolate out of a probably large number of stimuli the one stimulus or stimulus pattern to which the response is to be made, and conditioning is said to have been established only when the animal has begun to make the conditioned response more or less specifically to the conditioned stimulus. In instrumental learning, even more clearly than in conditioning, the animal is made to differentiate the manifold aspects of the situation. Of all the responses which may and do occur in the course of instrumental learning, only the 'correct' one is regularly and specifically reinforced. Inevitably, therefore, the experimenter is imposing a discrimination procedure, and, if the animal learns to make the 'correct' response, it may be said to have learned the discrimination. In contrast to these conditions, the procedures most commonly followed in experimental extinction do not require differentiation of the many activities which occur while the extinction is in progress. During the extinction of a conditioned response the animal is free to learn to stop responding to virtually any stimulus, since the unconditioned stimulus no longer differentially reinforces the response made to the conditioned stimulus. When an instrumental response is being extinguished the animal is no longer given the particular reward for making the 'correct' response, but neither is it given that reward for making other responses; accordingly, the animal is not required to differentiate not-responding with the previously 'correct' response from not-responding with any other response.

In developing the implications of the aforementioned difference between the procedures of positive and negative learning, it is necessary, at this point, to introduce two new assumptions. It must be assumed, first, that differentiations which are established in the course of positive learning are not wholly carried over to the corresponding extinction or negative learning; and, second, that the transfer of learning will depend largely on the degree to which the procedures

of learning require that differentiations be made.[3] (The second assumption is a corollary to the view, implied in the preceding paragraph, that learning is essentially the procedure by which the various stimulus and response patterns become differentiated.) From these assumptions and from the analysis of positive and negative learning procedures presented above, it follows that the transfer of extinction will tend normally to be relatively broad.

According to the hypothesis outlined in this paper, it should be possible to restrict the transfer of extinction, and so reduce spontaneous recovery, by applying procedures which cause the animal to differentiate the to-be-extinguished from not-to-be-extinguished activities. Such procedures are the basis of ordinary discrimination learning, in which one response is specifically reinforced and another is specifically not reinforced, and the data on recovery of the non-reinforced response should, therefore, be relevant to our hypothesis. In this connection Pavlov (**5**, p. 123) has concluded from his investigations of discrimination learning that the recovery of the extinguished response becomes less as "the differentiation [between a positive and a negative stimulus] becomes more firmly established." More recently, Hilgard and Humphreys (**1**) have found little spontaneous recovery of a response which had been extinguished in the process of discrimination learning. These observations provide some support for the hypothesis which is being presented in this paper; unfortunately, however, they do not permit a direct estimation of the effects of differentiation on spontaneous recovery, since in Pavlov's experiments the progressively diminished recovery may have resulted from the greater accumulation of non-reinforced trials, and in the experiment by Hilgard and Humphreys there is no measure of the amount of recovery which would have occurred if the differentiation had not been established.

The experiment to be reported in the succeeding pages was designed to provide a more nearly direct measure of the effect on spontaneous

[3] In its most moderate form, this assumption is frequently made, and implies for conditioning that the application of discrimination procedures will effectively limit stimulus and response generalizations. Spence (**6**) and Hull (**2**), for example, have worked out mathematical formulations for discrimination learning which describe the restriction of effective stimulus generalization as a consequence of the algebraic interaction of the generalization gradients corresponding to the positive and negative tendencies. A more extreme form of this assumption would propose that transfer in general, or stimulus generalization in particular, is initially complete, and becomes restricted only as differential reinforcements occur in accordance with the demands of past and present learning situations. In this connection, Lashley and Wade (**3**) have recently emphasized that a wide variety of stimuli is adequate to elicit the conditioned response during the early stages of learning, wherefrom these authors conclude that "the conditioned reaction is initially undifferentiated." For the purposes of the hypothesis being presented in this paper it is necessary merely to assume that the overall transfer which may result from some as yet unanalyzed combination of stimulus and response generalization will be most significantly determined by the degree to which differentiation is required in any given learning situation. This assumption and its relevance to the present hypothesis will be described and analyzed in greater detail in a paper to be published soon.

recovery of an extinction procedure which causes the animal to differentiate the to-be-extinguished activity from a not-to-be-extinguished activity. Such a procedure is calculated to limit the otherwise broad transfer of the extinction tendency and thus reduce the number of activities which have a potentiality for interfering with retention of the extinction. It is to be expected, therefore, that a response which has been extinguished by this procedure will undergo relatively little recovery by comparison with the extent to which the same response will recover when differentiations have not been established in the course of extinction.

Procedure and Apparatus

The following paragraph is a summary of the experimental procedure. Two groups of rats were trained to a performance criterion on each of two runways, A and B. One group, to be known as the double-extinction group, was extinguished to a performance criterion on Runway A by a procedure which involved non-reinforced trials on Runway A, irregularly interspersed with non-reinforced trials on Runway B. The other group, to be known as the differential-extinction group, was extinguished to the same performance criterion on Runway A by being given a series of non-reinforced trials on Runway A, irregularly interspersed with reinforced trials on Runway B. (The procedure for the latter group was intended to produce the greater differentiation of the runways, and to lead, therefore, to the lesser amount of spontaneous recovery.) Both groups were tested for spontaneous recovery on Runway A one hour after the extinction criterion had been met.

Runway A was a gray wooden box, eight feet in length, four in. in width, and five in. in height (inside dimensions). Coarse wire netting covered the floor of the runway. One and one-half feet down from one end of the box was a drop-shutter which could be lowered into the box so as to define a starting area in which the animal could be confined for a few seconds at the beginning of each trial. At the opposite end of the box from the starting area was the food dish. A gray canvas curtain, suspended from the wall of the runway box at a point just above the food dish, covered the food dish and its contents, and the food could not be seen or taken by the rat except as the curtain was pushed aside.

Runway B was a black, L-shaped box, seven and one-half feet in overall length, six in. in width, and four in. in height. The foot of the 'L' began at a distance of four and one-half feet from the start of the runway. The floor of the runway was of wood. A starting area, at one end of the runway box, could be separated from the rest of the runway by a drop-shutter. At the other end of the runway was a food compartment, one foot wide, one foot long, and four in. high. In one corner of this food compartment was the food dish, and between the food dish and the entrance to the food compartment was a partition which made it impossible for the food dish to be seen from any position behind a 'finish' line drawn across the food compartment about five inches from the end of the runway.

The specific procedures for the animals in both groups were as follows.

For seven days immediately preceding the beginning of the experiment the rat was fed each day an amount of food which constitutes about one-half to two-thirds of the daily caloric requirement. Throughout this seven-day period the animal was fed at the same time each day. The experimenter handled the animal for three or four min. before the food was given. In no case did the animal feed from the experimenter's hand. On the eighth day following the beginning of this regime the experiment proper was begun.

The first day of the experiment was taken up with procedures calculated to habituate the animal to the two runways. On each trial of this habituation the animal was placed in the runway at a point very near the food dish. When the animal had found and eaten a 0.08-gram pellet of food (dog-chow), it was removed from the runway. Sixty sec. intervened between trials, during which time the rat was kept in a waiting cage. There were five consecutive habituation trials on Runway A, followed by five consecutive trials on Runway B. On each trial a record was kept of the amount of time taken by the animal to find and eat the food.

On the second day of the experiment the rat was trained, in both runways, to run from the starting box to the food dish. At the beginning of each training trial the rat was placed in the starting box. Within five sec. thereafter, the shutter was raised, and the animal proceeded ultimately to make its way to the food dish where it found and consumed the 0.08-gram pellet of dog-chow. When the pellet of food had been eaten, the rat was removed to the waiting cage for one min., after which the next training trial was begun. On each trial the experimenter measured the running-time with a stop-watch. The watch was started as the shutter was raised; the watch was stopped, in the case of Runway A, when the rat touched the curtain which covered the food dish; in the case of Runway B, the watch was stopped when the rat crossed the 'finish' line. Training was given first on Runway A, and was continued on that runway until the rat had made three runs in any five trials in 10 sec. or less, or until the rat had been given a total of ten trials, whichever event occurred first. The animal was then trained on Runway B to the same criterion of performance or trials.

The third experimental day was taken up with retraining, extinction according to two different procedures, and then test for spontaneous recovery. The animal was first given a series of retraining trials on Runway A until it had run three trials in any five in 10 sec. or less, after which it was retrained to the same criterion on Runway B. Except for the fact that all animals were run until they had met the performance criterion, the procedure for retraining was identical with that which had been applied in the training session.

Extinction trials were begun one min. after the retraining on Runway B had been completed. In the double-extinction group all rats were given a series of non-reinforced trials on Runway A and non-reinforced trials on Runway B in the order: ABAABABBABABAABBABABBABABABBABAAB, etc. In the differential-extinction group the rats were given non-reinforced trials on Runway A and reinforced trials on Runway B in an order identical with the above. (Trials on Runways A and B were presented in this irregular order so as to decrease the probability that the rats in the differential-extinction group would learn to respond to the simple alternation of reinforced and non-reinforced trials.) These extinction trials were spaced at 30-sec. intervals, and the rat was always replaced in the waiting cage between trials. The extinction procedures were continued in every case until the rat had run two trials in any five on Runway A in 30 sec. or more. When the extinction criterion had been met, the animal was placed in the waiting cage for one hour. At the end of the one-hour period, each animal was given five consecutive non-reinforced trials on Runway A. As in the preceding extinction series, these trials were spaced at intervals of 30 sec., and the rat was replaced in the waiting cage between trials. The running-time on these five non-reinforced trials will serve as the measure of spontaneous recovery.

Of the 42 rats on which complete data were secured, 30 were hooded males from the Yale colony and 12 were male albinos from the Albino Farm, Redbank, New Jersey. One-half of the hooded rats served in the double-extinction condition, and one-half in the differential-extinction condition; similarly, one-half of the albinos served in each condition. All rats were between the ages of three and five months when used in the experiment.

Results

The values in Table I reflect the performance of both groups of animals on the last two[4] of the five trials which constituted the habituation procedure.

Inasmuch as the differences between the group means in Table I fall far short of statistical significance, neither of the groups may be said to have habituated more readily than the other.

In Table II are the data which describe the results of the initial and final stages of the two training sessions and provide a basis for determining whether the groups are equated with respect to performance during training.

[4] Meaningful measures of performance on the first three trials of habituation are unavailable in this study, inasmuch as it was sometimes necessary on those trials to improvise procedures for animals which were especially slow about finding the food.

TABLE I

MEANS AND STANDARD ERRORS REPRESENTING THE AMOUNTS OF TIME (IN SEC.) TAKEN TO FIND AND EAT THE FOOD ON THE LAST TWO TRIALS OF HABITUATION

	Trial 4		Trial 5	
	Mean	σ_m	Mean	σ_m
Runway A				
Double-extinction	54.26	11.92	32.04	8.09
Differential-extinction	41.38	10.80	37.13	8.11
Runway B				
Double-extinction	40.21	9.76	43.18	10.05
Differential-extinction	50.14	11.04	45.12	10.32

It will be concluded that the two groups do not differ in performance at the beginning and end of either training session, since the mean differences between the groups are inconsistent as to direction, and the most nearly significant of them is such that a difference as large or larger would have arisen 48 times in 100 by chance.

In the course of the first training session, six of the 21 rats in the double-extinction group and five of the 21 rats in the differential-extinction group met the performance criterion (three runs in five of 10 sec. or less) on Runway A in fewer than 10 trials; on Runway B that criterion was met in fewer than 10 trials by 11 animals in the double-extinction group and by 10 animals in the differential-extinction group. (Those animals which did not meet the performance criterion at the first session were given a total of 10 trials.) In the second training session all animals (including those which had met the criterion at the first session) were brought to the aforementioned criterion. This criterion was reached on Runway A in an average of 4.57 trials by the double-extinction group and in an

TABLE II

MEANS AND STANDARD ERRORS OF THE RUNNING-TIMES (IN SEC.) ON THE FIRST AND LAST TRIALS OF THE FIRST AND SECOND TRAINING SESSIONS

	First Training Session				Second Training Session			
	First Trial		Last Trial		First Trial		Last Trial	
	Mean	σ_m	Mean	σ_m	Mean	σ_m	Mean	σ_m
Runway A								
Double-extinction	36.38	6.30	9.95	1.51	21.05	6.04	4.67	0.52
Differential-extinction	40.81	7.49	11.62	1.86	20.71	6.11	4.90	0.44
Runway B								
Double-extinction	39.48	7.01	11.57	2.17	13.33	3.36	5.67	0.41
Differential-extinction	34.52	5.38	11.67	2.30	11.29	1.46	5.95	0.40

average of 4.86 trials by the differential-extinction group; on Runway B the double-extinction group required 4.62 trials and the differential-extinction group required 4.71 trials. Although the double-extinction group reached the criterion on both runways in the lesser number of trials, the differences between the groups in this regard are relatively small, and a statistical analysis reveals that differences as large or larger than those obtained would have arisen by chance more than 69 times in 100 (Runway A) and 88 times in 100 (Runway B).

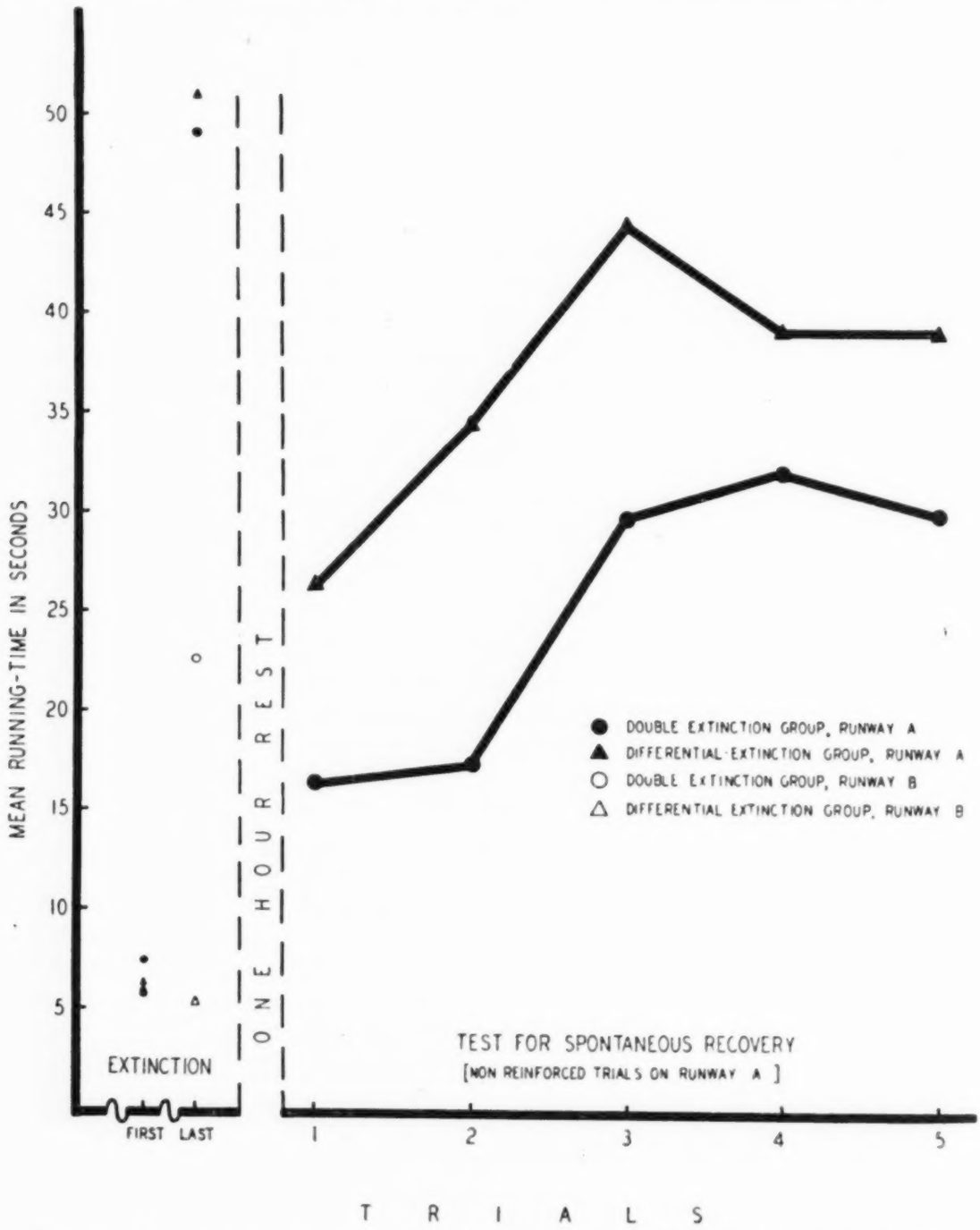


FIG. 1. Mean running-times in sec. on the first and last extinction trials and on the five trials of test for spontaneous recovery

When the training had been completed, the two groups of animals were extinguished to a common performance criterion on Runway A by procedures which were calculated to produce a difference in the degree to which the groups differentiate Runways A and B. The one group (double-extinction) was given a series of non-reinforced trials on Runway A, irregularly interspersed with non-reinforced trials on Runway B. The other group (differential-extinction) received non-reinforced trials on Runway A, irregularly interspersed with reinforced trials on Runway B.

The mean running-times on the first trial of extinction, as shown in Fig. 1 and Table III, indicate that the two groups were very nearly equated in performance on both runways at the beginning of the extinction series. On the first trial of extinction the more significant of the two mean differences between the groups is such that a difference as large or larger would have arisen more than 29 times in 100 by chance. A comparison of the mean running-times for the last trial of extinction shows that the groups reached nearly equivalent final levels of extinction on Runway A. Mean differences as large or larger than the one obtained on the last trial of extinction would have arisen by chance more than 62 times in 100.

TABLE III

MEANS AND STANDARD ERRORS OF THE RUNNING-TIMES (IN SEC.) ON THE FIRST AND LAST EXTINCTION TRIALS AND ON THE FIVE TRIALS OF TEST FOR SPONTANEOUS RECOVERY

		Extinction Trials		Recovery Trials				
		First	Last	First	Second	Third	Fourth	Fifth
Runway A								
Double-extinction	Mean	7.38	49.24	16.43	17.33	29.67	32.00	29.83
	σ_m	1.17	2.45	3.15	3.54	4.40	4.78	4.63
Differential-extinction	Mean	5.81	50.95	26.24	34.24	44.33	39.05	39.00
	σ_m	0.91	2.60	4.57	4.84	4.55	4.44	3.71
Runway B								
Double-extinction	Mean	5.76	22.62					
	σ_m	0.41	4.83					
Differential-extinction	Mean	6.24	5.29					
	σ_m	0.47	0.81					

To determine whether the extinction procedures of this experiment did produce more and less differentiation of the runways, the groups may be compared with regard to their performance on both runways at the end of extinction. In this connection it has already been seen that the two groups reached equivalent levels of extinction

on Runway A. A further examination of the means in Fig. 1 and Table III shows that the final running-time on B is considerably lower for the differential-extinction group than for the double-extinction group. The difference between the groups in final running-times to B is statistically significant below the level of .01. It will be concluded, therefore, that the differential-extinction group has the significantly greater difference in running-time between Runways A and B, and may thus be said to have achieved the greater differentiation of the runways.

In the double-extinction group the final running-time to B is lower than the final running-time to A, even though the procedure for this group involved non-reinforced trials on both runways. One possible reason for the lower running-time to B in this group is to be found in the fact that the extinction criterion was applied only to Runway A. This circumstance guarantees that the final running-time to A shall equal or exceed the criterion (30 sec.), whereas no such provision is made for the running-times on Runway B. A second possible reason derives from the fact that there was, on the average, one more trial of extinction on A than on B. (This was a consequence of the order in which the trials were presented. See procedure.)

In the differential-extinction group the application of the extinction criterion to Runway A may have operated, as in the double-extinction group, to reduce the running-time to B by comparison with A. On the other hand, the greater number of trials on A should have had the effect in the differential-extinction group of increasing the running-time to B beyond what it would have been if the number of reinforced trials on B had been equal to the number of non-reinforced trials on A. In any event, neither of these aspects of the procedure could have produced that difference between the groups (in final running-time to B) which was taken as evidence for the conclusion that the differential-extinction group achieved the greater differentiation of the runways. That difference may be presumed to have been caused by the fact that trials on B were reinforced for the differential-extinction group and not reinforced for the double-extinction group.

According to the considerations outlined in the Introduction, differentiation of the runways should limit the otherwise broad transfer of the extinctive tendency and thus reduce spontaneous recovery by limiting the number of interim activities which can interfere with the retention of the extinction. Since the differential-extinction group achieved the greater differentiation of the runways, it is to be expected that there will be less spontaneous recovery in this group than in the double-extinction group. Confirmation of this expectation is to be had by comparing those mean running-times in Fig. 1 which are bracketed within the heading 'Test for spontaneous recovery.' These means illustrate the effects on spontaneous recovery of the two extinction procedures which were applied in the experiment, and a comparison of them reveals that the animals in

the differential-extinction group have the longer running-times (hence less spontaneous recovery) on each of the five tests for recovery. The probabilities that differences in mean running-time as large or larger than those obtained would have arisen by chance are $<.08$, $<.01$, $<.02$, $<.30$, and $<.11$ for the first, second, third, fourth, and fifth trials, respectively.

It will be noticed that the difference between the two groups is relatively larger and more nearly significant on the second and third trials of the recovery test than on the first trial. This is to be accounted for, perhaps, on the grounds that disinhibitory effects, resulting from the rather sudden re-introduction of the animal into the experimental situation, would be most likely to operate and to override other influences on the first trial following the rest period. On the fourth and fifth trials, also, the differences are smaller and less nearly significant than on the second and third trials. This is understandable on the assumption that the continued application of non-reinforced test trials on Runway A will tend to bring both groups to the same ultimate level of extinction and so reduce the difference between the groups.

It was to be expected that the double-extinction group would have required fewer trials to reach the extinction criterion, since the transfer of the extinctive tendency from B should have facilitated extinction on A for the double-extinction group, whereas transfer of the reinforced running on B should have retarded extinction on A for the differential-extinction group. In fact, the criterion of extinction was reached by the double-extinction and differential-extinction groups in averages of 10.28 and 11.00 trials,[5] respectively. Although this difference is in the expected direction, it is hardly significant, inasmuch as a difference as large or larger would have arisen 69 times in 100 by chance. In this connection, however, it is noteworthy that a retabulation and further analysis of the extinction data reveals that the double-extinction group required the consistently smaller number of trials to reach extinction criteria less strict than the criterion used in this experiment. The less strict criteria are all in terms of two runs in any successive five on Runway A in x seconds (criterion time) or more, and the mean number of trials taken to reach these several criteria are as follows, the mean for the double-extinction group being given first in every case: criterion of 25 sec., 8.76 and 9.62; criterion of 20 sec., 7.38 and 8.57; criterion of 15 sec., 6.38 and 6.95; and criterion of 10 sec., 4.67 and 4.86.

There is the possibility that the lesser spontaneous recovery of the differential-extinction group may be accounted for in terms of the greater number of trials taken by this group to reach the extinction criterion. To provide a basis for evaluating this possibility, there are given in Table IV the correlations between number of trials to extinguish and running-times on each of the five trials of the recovery test.

[5] These values represent the number of trials on Runway A and do not include the interspersed trials on Runway B.

TABLE IV

PRODUCT-MOMENT CORRELATIONS BETWEEN NUMBER OF TRIALS TO EXTINGUISH AND RUNNING-TIME ON EACH OF THE FIVE TESTS FOR RECOVERY, TOGETHER WITH THE PROBABILITIES THAT CORRELATIONS AS LARGE OR LARGER WOULD HAVE ARISEN BY CHANCE

(Double-extinction)	*r*	*P**	(Differential-extinction)	*r*	*P*
1st recovery trial	+.27	<.25	1st recovery trial	+.13	<.62
2nd recovery trial	−.28	<.25	2nd recovery trial	+.04	
3rd recovery trial	−.40	<.11	3rd recovery trial	−.02	
4th recovery trial	−.21	<.38	4th recovery trial	−.16	<.49
5th recovery trial	−.16	<.49	5th recovery trial	−.40	<.11

* '*P*' refers to the probability that a positive or negative correlation as large or larger than the one obtained would have arisen by chance if the true correlation were zero.

A significantly high positive correlation (between number of trials to extinguish and running-time on the recovery test) would support the assumption that the relatively lesser spontaneous recovery (longer running-time) of the differential-extinction group may be attributed, in some measure at least, to the fact that this group took the greater number of trials to extinguish. The obtained correlations (Table IV) are not significantly different from zero, and a majority of them are negative. It may be concluded, therefore, that the relatively greater number of trials taken by the differential-extinction group does not account for the fact that this group has the relatively lower amount of spontaneous recovery.

SUMMARY

It had been proposed in an earlier paper that spontaneous recovery can be accounted for, in part at least, on the assumption that the relatively broad transfer of extinction tendencies sets up a large number of activities which can and do interfere with the retention of the extinction. In the paper being summarized here it is suggested that the normally broad transfer of extinction is to be attributed to the circumstance that the typical procedures of extinction (by comparison with the procedures of positive learning) require relatively little differentiation of the stimulus or response aspects of the learning situation. These suggestions imply that spontaneous recovery should be reduced in amount as a result of the application of extinction procedures which limit the transfer of the extinctive tendency by causing the animal to differentiate the to-be-extinguished from not-to-be-extinguished activities. In order to provide evidence relevant to this implication, the following experiment was carried out.

Two groups of rats were trained by the same procedure to a criterion of performance on Runways A and B. One group (double-extinction) was extinguished to a performance criterion on Runway A by being given non-reinforced trials on A, irregularly interspersed

with non-reinforced trials on B. The other group (differential-extinction) was extinguished to the same criterion on A according to a procedure which involved non-reinforced trials on A, irregularly interspersed with reinforced trials on B. The latter procedure produced in the differential-extinction group the relatively greater differentiation of the runways (as indicated by performance measures). Measures of spontaneous recovery taken one hour after the end of extinction reveal that the differential-extinction group had the lesser amount of spontaneous recovery.

(Manuscript received December 29, 1947)

References

1. Hilgard, E. R., & Humphreys, L. G. The retention of conditioned discrimination in man. *J. gen. Psychol.*, 1938, **19**, 111–125.
2. Hull, C. L. The problem of stimulus equivalence in behavior theory. *Psychol. Rev.*, 1939, **46**, 9–30.
3. Lashley, K. S., & Wade, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, **53**, 72–87.
4. Liberman, A. M. The effect of interpolated activity on spontaneous recovery from experimental extinction. *J. exp. Psychol.*, 1944, **34**, 282–301.
5. Pavlov, I. P. *Conditioned reflexes.* (Trans. by C. V. Anrep.) London: Oxford University Press, 1927.
6. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 130–144.

DISCRIMINATION OF THE DURATION OF A SOUND

BY FRANKLIN M. HENRY

University of California, Berkeley *

Introduction

Gilliland, Hofeld and Eckstrand (**5**), reviewing the literature on time perception since 1933, emphasize that "inconsistencies have been obtained in determining whether Weber's law holds in time estimation." In one study by Gilliland (**4**) the law seemed to be verified for a duration range of 4–27 sec., although a later study by Gilliland and Humphreys (**6**) found the percent errors in discrimination were increased for the short time intervals in the range 9–180 sec.

There are only a few modern studies of the difference limen for the duration of a sound, and these offer only incomplete descriptions of the stimuli. Zwirner (**14**) in 1938 stated that the limens follow the Weber-Fechner law. Stott (**12**) in a study of time-order errors, published data in 1935 from which difference limens may be calculated—the Weber ratio (for 75 percent correct judgments) for a 1000 cycle tone of unspecified intensity is 0.145 for a duration of 200 msec., drops to 0.115 at 400 msec., remains at this value for intermediate durations, then tends to rise again with durations longer than one sec. His paper also discusses the necessity for using electronic equipment to achieve proper control of acoustic stimuli.

The present investigation is concerned with the nature of the difference limens for duration of very brief sounds ranging from about 30 msec. to nearly ½ sec. in length, with some exploration of the effects on this limen of variations in stimulus quality and intensity.

Apparatus

Stott ran into some difficulties with his stimulus timer at short durations, even though he dealt with durations longer than those used in the present investigation. This pointed to the necessity for careful apparatus design. After several abortive models were tested, a cam-operated timer was constructed that incorporates the following features: (a) continuously adjustable stimulus durations; (b) independent circuits for standard and experimental durations; (c) continuously adjustable increment that can be added or deleted quickly.

The function of the apparatus is to control the duration of the elements of a series of tones so that (a) a standard tone of specified duration is repeated regularly, or (b) the standard tone alternates with an experimental tone that differs from the standard by a specified increment of duration. This is accomplished by interposing a set of contacts between a generator of audio frequency electric currents and a telephone that converts the electric currents into sound waves.

* From the Division of Medical Physics and the Department of Physical Education.

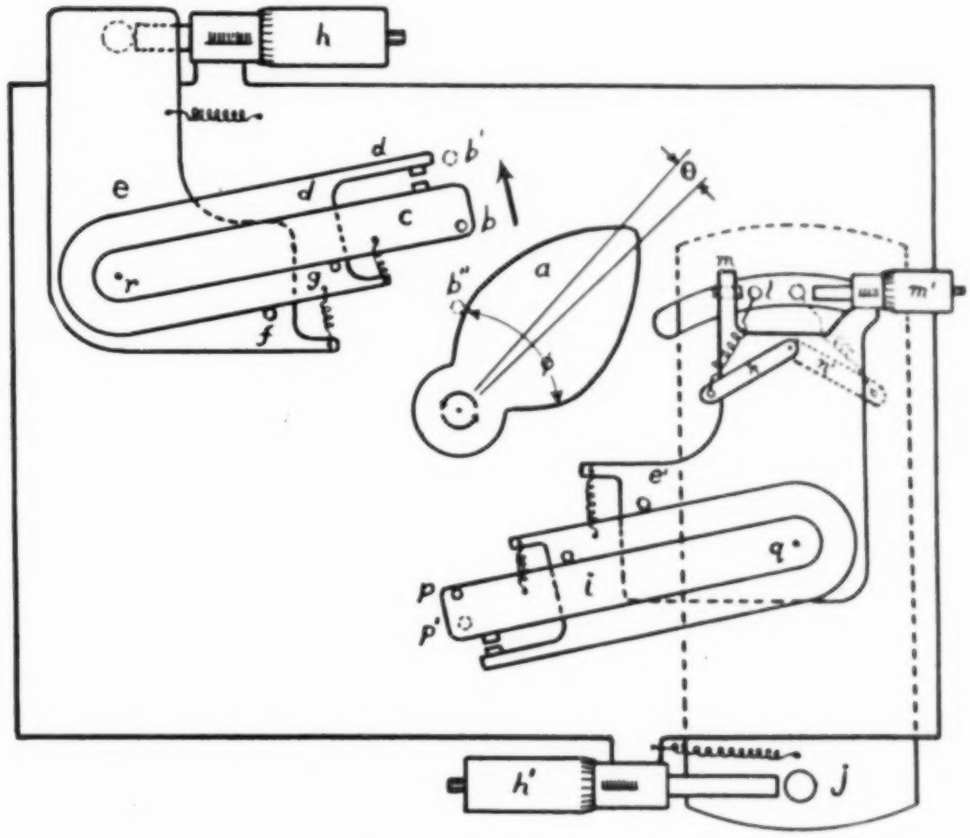


FIG. 1. Stimulus timer

The apparatus is shown diagrammatically in Fig. 1. Duration is controlled by micrometer *h*; when this is set at zero, contacts *c* and *d* just fail to close as the cam a rotates. As the micrometer is opened, the contacts close for a period of time proportional to the micrometer setting. Micrometer *h'* (at the bottom of the figure) controls another set of contacts; provision is made for shortening the duration of these contacts by an increment controlled by micrometer *m'* if the lever *n* is snapped over to *n'*.

A cam of modified cardioid shape is rotated by a constant speed motor at the rate of 0.83 rps. This cam shape yields a substantially linear relation between stimulus duration and micrometer reading; the advantage of this is that a single calibration curve gives the increment for all durations. The cam-follower *b* is pivoted at *r*, so that as the cam turns it slides against the roller-pin *b* and forces it up to a new position *b'* at the peak of the lift. The first upward movement of the follower has caused contacts *c* and *d* to close; thereafter the auxiliary cam-follower *d* moves with the main cam-follower, the contacts remaining closed. After the peak of the cam has passed, the followers are lowered until *d* strikes the limiting pin *f* which stops it. Further rotation lowers *c*; since *d* is stopped the contacts open. The resting position of the main follower is fixed by cam-stops *g* and *f*; this position is adjustable since stop *f* is mounted on an L-shaped piece *e* pivoted at the common center *r*, and controlled by the micrometer *h* which is mounted on the base plate of the apparatus.

The other cam-follower-contact assembly (controlled by micrometer *h'*) operates in much the same fashion, except that its adjustment is indirect through a rectangular piece of metal mounted under the base-plate and pivoted at the common point *q*. A heavy pin at *j* serves as the anvil for micrometer *h'*; another pin at *l* passes through a slot in the base-plate and is held against the L piece at either *m* (zero increment) or at the end of micrometer *m'* (increment in).

Naturally the various springs are graduated in stiffness so that the various parts of the apparatus operate in proper sequence. They are packed with felt to avoid parasitic vibrations. The cam-follower arms are made from bakelite to simplify the problem of insulation. All dimensions are laid out precisely so that the three micrometers have equivalent effects on stimulus duration.

Pivot bearings are carefully lapped in so there is no free play to cause inaccuracy. The construction is relatively massive.[1]

Additional contacts (not shown) are mounted on the lower left corner of the base plate and opened by the tip of the cam. These are in series with the motor and shunted by a push button. Pushing the botton starts the motor, which then runs through one pair of stimuli and stops automatically when this type of operation is desired.

The timer contacts (in parallel) connect the anode potential to an amplifier through a resistor of 10,000 ohms that is by-passed by a half-microfarad capacitor. This has the same function as the grid control circuit used by Stott, insuring that the earphone current produced by the pure tone electronic oscillator is turned on and off without discernible clicks. Oscilligraphic measurements yielded a half-time of 2.5 msec. for the rise and decay of the current. Interposed between the oscillator and earphone is a 'T' type attenuator pad of conventional design that controls the sound intensity in steps of one db.

The timer was calibrated originally with a Dunlap chronoscope. It was discovered that this instrument gave erratic results for durations of 30 msec., so a recalibration was made with an oscillograph.

Method

The *S*s were tested individually in a sound-proofed room and were in telephonic communication with the *E*, who controlled the apparatus in an anteroom. First a series of sounds was demonstrated in which the standard alternated with a sound differing in duration by the increment being tested; then came a demonstration series in which all sounds had equal standard length followed by another demonstration with the increment as in the first series. Following this 'known,' the *S* was given an 'unknown' series and was asked to judge whether the sounds were constant in duration or alternately differed in duration. When a judgment was reached, another 'unknown' was presented, and so on until 10 unknowns had been judged, with 'increment' or 'no increment' occurring in chance order. The increment was constant in any such set of 10 judgments. After a brief rest, another set of knowns with a different increment and duration was presented, followed by 10 judgments as before. This was continued for an hour each day. For each *S*, a balanced order of increments and durations was followed so that practice and fatigue effects would be minimized. Eventually every *S* produced 50 judgments for each duration and increment. Twice this amount of data was secured in the case of Experiment 1.

No attempt was made to secure complete psychophysical curves; it was thought to be more economical to concentrate on increments that the three or four day training period indicated were fairly close to the limen (defined as 75 percent correct judgments). Absolute intensive thresholds for half-second tones were determined during the training period. The sound intensities used in the experiment represent db above these threshold values.

The writer is indebted to the Psychology Department, University of California, for providing the test-room facilities.

Results

Experiment 1.—The first experiment was concerned with the constancy of the Weber ratio at various stimulus durations. A tone of medium pitch (500 cycles) and moderate loudness (50 db above threshold) was used. The durations were of odd lengths because the collection of data started before the timer was calibrated; they approximate a geometric series with a 15 fold range from shortest to longest.

The results from testing seven junior and senior psychology students (three women and four men) are shown in Table I. It is

[1] Photographs are available to anyone who is interested in a more complete description of the apparatus.

TABLE I

MEAN WEBER RATIO $\Delta T/T$ FOR SEVEN *Ss* TESTED AT SEVEN DURATIONS

The stimulus was a 500 cycle tone 50 db above threshold. (The critical ratio *t* and intercorrelation *r* between adjacent columns is also given.)

Duration (msec.) =	32		47		77		110		175		277		480
Weber ratio =	.281		.203		.208		.196		.188		.172		.143
SD =	.067		.034		.052		.061		.026		.025		.022
t ratio =		8.2		0.4		1.4		0.4		0.9		7.1	
r =		.89		.82		.94		.57		.86		.92	

apparent from the size of the critical ratios that significant changes occur in the Weber ratio. Particularly striking is the large increase between 32 and 47 msec. duration (Fig. 2). Aside from this sharp increase at the shortest time used, there is a slight tendency for the Weber ratio to decrease more or less linearly with increased duration over the stimulus range studied. A change in duration of 50 to 100 msec. does not result in a statistically significant drop, but a larger change, such as the 230 msec. jump from the 47 msec. duration to the 277 msec. duration significantly lowers the ratio (CR = 4.1); likewise, the 203 msec. jump between the two longest durations has a significant effect, as may be seen in the table.

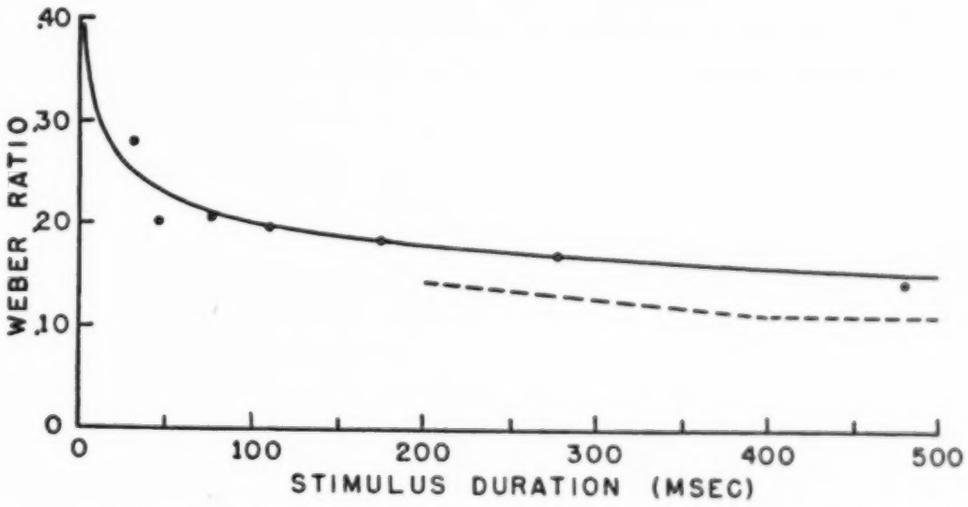


FIG. 2. Size of the Weber ratio for time discrimination as a function of the duration of the tone. (The dotted line shows the Stott data for comparison.)

The Weber law is usually conceded to hold for the middle range of stimulus intensities for a variety of sense modes and types of discrimination; it is also commonly known that the ratio quite universally rises at low stimulus magnitudes. To account for this fact, Guilford (7) has proposed a power law:

$$\Delta R = kR^N.$$

From this equation, we may write:

$$\Delta T/T = kT^{N-1},$$

which predicts a rising Weber ratio at the short durations as shown by the smooth curve in Fig. 2.

The rate at which the ratio is lowered—about .027 ratio points per 200-msec. increase in duration for the flat part of the curve—is quite similar to the amount of decrease calculated from the Stott data cited in the introduction. On the other hand, the absolute agreement between the two experiments in the region where they can be compared is not so good; the interpolated ratio of .184 is probably significantly higher than the Stott ratio of .145 at 200 msec. and the interpolated ratio of .154 also seems considerably higher than the Stott figure of .115 at 400 msec. (Critical ratios cannot be calculated since the Stott data do not include measures of variability.) There is probably some overlap—one *S* in the present experiment reached the Stott mean—but the difference may well be due to other factors than sampling error. The intensity of his sounds were not given. Their frequency was higher. Both experiments used a variation of the constant method, but the methods were not identical. Another point to be considered is the nature of the stimulus—the Stott apparatus probably gave a much sharper beginning and end to the sound which would probably lead to somewhat lower thresholds. Also, there is always the possibility of calibration errors in experiments on short time durations.

TABLE II

MEAN WEBER RATIO $\Delta T/T$ AS A FUNCTION OF STIMULUS INTENSITY (IN DB ABOVE THRESHOLD)

Results are given for three durations.

Intensity	Duration			Average
	47 msec.	77 msec.	277 msec.	
20 db	.204	.198	.149	.184
40 db	.153	.184	.137	.158
60 db	.166	.179	.147	.164
80 db	.157	.184	.139	.160
Average	.170	.186	.143	.166

Experiment 2.—In this experiment, the stimulus intensities ranged from 20 db (a relatively faint sound about as loud as an ordinary whisper) to 80 db (quite loud, and somewhat uncomfortable). The frequency was 500 cycles as in the first experiment. Data were secured on five *S*s at the 47 msec. duration, four *S*s at 77 msec., and five at 277 msec. Two of the *S*s who had served in Experiment 1 participated in all three parts of this experiment; two others did both 47 and 277 msec. tests and one did both 47 and 77 msec. tests.

The data are shown in Table II. Variation of stimulus intensity appears to have little effect except for a tendency for a somewhat higher Weber ratio for the 20 db sound. Considering the 47 msec. sound, this difference is significant at the one percent level of confidence when the 20 db stimulus is compared with the 40 db; significant at the seven percent level when 20 and 60 db are compared; significant at the two percent level when 20 and 80 db are compared. At 77 msec. the same tendency for a higher ratio at low intensity is apparent, but the change is not statistically dependable. When the data at 47 and 77 msec. are averaged (which is a reasonable procedure since these durations are close together as compared with 277 msec.), the 20 db ratio is significantly higher than the ratio obtained by averaging the 40, 60, and 80 db data—in fact, the difference is significant at the 0.3 percent level of confidence. In contrast with this finding, the 277 msec. data fail to show an 'intensity' effect.

Comparing the Weber ratio at short and long durations, there is confirmation of the trend for a smaller ratio at the longer duration as found in Experiment 1. The effect is accentuated by the 'low intensity' effect, but is present at the higher intensities also. Averaging the 40, 60, and 80 db data yields a ratio of .159 at 47 msec. and .182 at 77 msec. compared with the significantly lower ratio of .141 at .277 msec. On the average the ratio is somewhat lower and closer to the Stott data in this experiment—.166 compared with .194 for the average of the same durations in Experiment 1 and .155 for the Stott experiment.

TABLE III

INDIVIDUAL WEBER RATIOS FOR 77-MSEC. TONES OF DIFFERENT PITCH

The stimulus intensity was 50 db above threshold.

Subject	Pitch (cycles per sec.)				
	125	250	500	1000	2000
S	.290	.208	.143	.130	.158
L	.260	.195	.169	.156	.117
F	.335	.221	.169	.182	.208
Average	.295	.208	.160	.156	.161

Experiment 3.—Data on the effect of stimulus qualtiy were obtained from three Ss. The duration of the constant sound was 77 msec.; the intensity was 50 db above threshold in all cases. Variation in quality was secured by using tones of different pitch. Table III gives the results. While the data are too few to justify the calculation of critical ratios, it does appear that there is a tendency for the Weber ratio to be highest for tones of low pitch.

Discussion of Results

Békésy (1) in 1929 experimentally equated the loudness of tones of variable duration and constant intensity with a standard tone of constant duration and variable intensity. For tones shorter than the standard his data, and also the data of more recent investigators of the problem (2, 3, 10, 11), can be shown to fit a Fechner Law type of equation having the form:

$$db = k \log T + C,$$

where db is the equivalent loudness of the variable sound having a duration T, and a physical intensity I that determines the constants k and c by simple linear equations. This formula implies that stimulus intensity and duration are integrated by the hearing mechanism, so that a change in either intensity or duration results in a corresponding alteration in perceived stimulus magnitude.

The duration limens from Experiment 1 may be converted into equivalent intensity limens by the above formula.[2] The results of the calculations, using a velocity constant that gives 10 db change in loudness for a tenfold change in duration, are shown in Table IV.

TABLE IV

Calculated Equivalent Loudness of the Sounds Used in Experiment 1 and Difference Limens for Duration Calculated as db Equivalent

Differential intensity limens for continuous tones from Riesz (11) are shown for comparison.

Duration (msec.) =	32	47	77	110	175	277	480
Loudness (db) =	38.2	39.0	42.0	43.6	45.6	47.6	50.0
Duration limen (db) =	1.08	.80	.82	.77	.75	.69	.58
Riesz limen (db) =	.83	.81	.74	.70	.66	.63	.59

For comparison, the directly determined differential intensity limens for 12 Ss reported by Riesz (11) for various absolute intensity magnitudes are also shown. The agreement between the two sets of data is very good, particularly when allowance is made for the small number of individuals tested, the obviously large differences in method, and the fact that the curve constants for the calculations were derived from the data of other investigators.

[2] The data of Békésy and the other investigators referred to above can also be fitted reasonably well by a power function of the form $db = at^k$; however, the classical logarithmic function seemed more appropriate for the present purpose.

Another formula, relating loudness to tonal duration and stimulus intensity by the function $db \times T = k$, has been proposed by Lifshitz (9). It fails to fit the other data, and also gives peculiar results in the present study. Loudness of the shortest tone (of the present experiment) calculated by this function is very close to the absolute threshold, although the sound was actually heard with ease and subjectively seemed to have a loudness not more than 15 or 20 db below the longest tone. The calculated Weber ratio for differential intensity based on the observed duration discrimination is tremendous—several hundred percent—at the longest duration, when it should be smallest. At the short durations it becomes smaller than the Riesz limens. This result is unreasonable in terms of the formula itself, which assumes that physical stimulus intensity in db is directly equivalent to linear time units of duration.

The similarity between the equivalent intensity limens (calculated from time discrimination) and true intensity limens is a potent argument for believing that discrimination of sound duration (for short tones) is fundamentally determined by a summation of the number of neurophysiological 'elements' available for the discrimination at any particular level of duration in intensity, as argued by Holway and Hurvich (**8**) for the general case of differential sensitivity. It may be noted that the tendency for filled intervals to be perceived as longer than unfilled intervals is consistent with this idea. Such a theory does not necessarily preclude the action of other mechanisms at longer durations.

The integrated intensity theory, coupled with the observed tendency for a relative increase in the difference limen at small stimulus magnitudes, predicts an 'intensity effect' in the direction of a larger Weber ratio as stimulus intensity is lowered at constant duration. Since decreasing either duration or intensity lowers the integrated stimulus magnitude, the effect should be greatest at short durations and low intensities. It will be recalled that the data of Experiment 2 (Table II) yield statistical evidence that this is so, although a larger effect might have been anticipated. (Fewer *S*s and fewer judgments per *S* were available in Experiment 2.)

The consequence of altering the frequency of the tone also appears to have been predicted by the integrated intensity theory. The Riesz data show increasingly larger difference limens for intensity as the frequency is lowered below 500 cycles, with very little change for frequencies between 500 and 2000 cycles. His frequency effect is slight at high intensities but very large at low intensities. With decreasing phenomenal intensity as in the present experiment, the observed frequency effect below 500 cycles would be anticipated. It may also be mentioned that Békésy (**1**) found a slower build-up of loudness with increasing time at low frequencies, which is consistent with the experimental findings if his data are interpreted by the integrated intensity theory.

Evidence supporting an integrated intensity theory of the differential intensity limen for sound has been presented by Upton and Holway (**13**). Their stimuli were of constant intensity as in Experiment 1 of the present study, and $\Delta I/I$ for their single *S* can be shown to exhibit the same type of relation with T that was found between $\Delta T/T$ and T. The curve constants are of course different, since their durations were much longer (1–120 sec.). Their experiment does not show the intensity effect that is clearly apparent in the Riesz data. Garner and Miller (**2**), studying short durations (20–750 msec.) in two *S*s, also found that intensity sensitivity involves temporal integration of the sound stimulus. Since they compared a continuous

tone of constant intensity with an experimental tone of variable intensity and short duration, their data are not directly comparable to the present experiment. However, their investigation does show that $\Delta I/I$ varies with intensity, frequency and duration in the same general way that $\Delta T/T$ varies in the present experiment. In a later study (3) they find that this temporal integration also occurs for threshold stimuli. All of these data agree in showing a change of the order of 10 db for a tenfold change in duration, which is the rate used in the present study.[3]

Summary and Conclusions

Limens for discrimination of the duration of a sound ranging from 32 to 480 msec. were determined for 11 *S*s. Stimulus intensity was varied from 20 to 80 db; pitch was varied from 125 to 2000 cycles. Somewhat more than 20,000 individual judgments were obtained.

The average *S* could discriminate approximately 14 percent change in the duration of a 500 cycle tone of moderate intensity.

Discrimination was only about half this good for the shortest duration studied; it was also poorer for faint sounds and low-pitched sounds.

The effects of stimulus alteration were consistent with a theory that presumes an integration of stimulus intensity with respect to duration. Perception of differences in the duration of sound stimuli shorter than a half-second can be considered as fundamentally a discrimination of intensity differences.

(Manuscript received October 24, 1947)

References

1. Békésy, G. v. Zur Theorie des Hörens. *Physik. Z.*, 1929, **30**, 115–125 and 721–745.
2. Garner, W. R., & Miller, G. A. Differential sensitivity to intensity as a function of the duration of the comparison tone. *J. exp. Psychol.*, 1944, **34**, 450–463.
3. Garner, W. R., & Miller, G. A. The masked threshold of a pure tone as a function of duration. *J. exp. Psychol.*, 1947, **37**, 293–303.
4. Gilliland, A. R. Some factors in estimating short time intervals. *J. exp. Psychol.*, 1940, **27** 243–255.
5. Gilliland, A. R., Hofeld, J., & Eckstrand, G. Studies in time perception. *Psychol. Bull.*, 1946, **43**, 162–176.
6. Gilliland, A. R., & Humphreys, D. W. Age, sex, method, and interval as variable in time estimation. *J. genet. Psychol.*, 1943, **63**, 123–130.
7. Guilford, J. P. A generalized psychophysical law. *Psychol. Rev.*, 1932, **39**, 73–85.
8. Holway, A. H., & Hurvich, L. M. On the discrimination of minimal differences in weights: I. A theory of differential sensitivity. *J. Psychol.*, 1937, **4**, 309–332.

[3] It should be noted the Békésy (1, p. 118 and p. 744) reports less change in loudness at higher sound intensities, whereas the Munson data (10) seem to show a greater change.

9. Lifshitz, S. Apparent duration of sound perception and musical optimum reverberation. *J. acoust. Soc. Amer.*, 1936, **7**, 213–221.
10. Munson, W. A. The growth of auditory sensation. *J. acoust. Soc. Amer.*, 1947, **19**, 584–591.
11. Riesz, R. R. Differential intensity sensitivity of the ear for pure tones. *Physic. Rev.*, 1928, **31**, 867–875.
12. Stott, L. H. Time-order errors in the discrimination of short tonal durations. *J. exp. Psychol.*, 1935, **18**, 741–766.
13. Upton, M., & Holway, A. H. On the psychophysics of hearing. I. Monaural differential sensitivity and exposure time. *Proc. nat. Acad. Sci., Wash.*, 1937, **23**, 29–32.
14. Zwirner, E. Phonologische und Phonometrische Probleme der Quantität. *Proc. 3rd int. Cong. phon. Sci. Ghent.*, 1938, 58–65.

RESPONSES TO 'MEANINGFUL' AND 'MEANINGLESS' SOUNDS

BY R. C. DAVIS

Indiana University

The effect of a stimulus upon an individual is studied either by securing the *S*'s judgment, to get the so-called sensory effects, or by recording a muscular activity or adjustment, to get the so-called motor effect. The relation between the two has often been discussed in a speculative way, but without much direct evidence. Since it has become possible to measure muscular responses more exactly, the problem may be attacked experimentally. Using *S*'s judgments many relations between 'sensory effects' and stimulus characteristics have of course been ascertained. The relation of 'motor responses' to these same variables remains largely unknown.

In undertaking to study them an important difference between the normal 'sensory' and 'motor' procedures needs to be regarded. In most experiments, when the *S* is asked to judge stimuli, he is instructed beforehand that he is to make the judgment or report. That is, a stimulus is given a 'meaning,' in that it calls forth a prescribed response. It is usually taken for granted then that the same sort of event takes place whether it has this 'meaning' or not. But possibly this is not true. It is possible that telling the *S* to report changes the whole process instead of just adding something to what would have been there anyway. Of course it is well known that if an *S* is later asked to report on something he has not been told to look for he will do quite poorly. Is this because it is too late for the *S* to attach the final step, the report, to the process, or because the process itself occurs on a very reduced scale? So long as the *S*'s report is the criterion of the process the question is naturally unanswerable. But if an involuntary muscular response is the criterion we can look for an answer. Then due allowance for the effect of instructions can be made in the comparison of 'sensory' and 'motor' effects. So the purpose of the present experiment is to investigate the effect of instructions to respond, using involuntary muscular reactions as a criterion.

An earlier experiment (1) showed what effects are produced on the muscular reactions by a preliminary 'voluntary' tension. The muscular responses are thereby enlarged and also concentrated at

the seat of the tension. Instructions to respond to the coming stimulus might work in the same fashion, if they produce a continuously raised tension somewhere. That they do so was a conclusion of study of set (3). It was therefore hoped that the present experiment if it found an effect of instructions might also show how it comes about.

Procedure

The method of experimentation was quite similar to that used before (1). *S* lay on his back on a cot. After five minutes of rest, 12 tones were given to him at 2-min. intervals (approximate). The general instructions to the *S* were to remain just as relaxed as possible. Sometimes, when the *S* did not relax well at first, he was given more specific instructions about relaxing, particularly his arms. For half the stimuli the *S* was under instructions to do nothing. For the other half he was instructed to do nothing until the end of the tone, then to press a button on which his right forefinger always rested. There were 24 *Ss* (all students and all right handed). Twelve of them had their instructions to respond for the first six stimuli; the other 12 were told nothing about responding until the last six stimuli were about to come.

The stimuli were three intensities of a 925 cycle tone, 70, 90, and 100 db. The *S* heard two sounds of each intensity while he was to respond, and two of each while he was not. Six different orders of stimulus presentations were used for *Ss* who 'responded' to the first six stimuli, two *Ss* having each order, and the same six orders were used for *Ss* who 'responded' to the latter six, two *Ss* having each order. So in all we might say there were 12 orders of presentation. The mechanism for controlling the time and duration of stimuli was that described in a previous note (2).

As in the earlier experiment (1), 'active' electrodes were located over the extensor muscles in the forearm, in a corresponding position for each *S*, as nearly as could be done. The reference electrodes were on the upper arms. The amplifiers were the same sort as used before; four stages of push-pull design, with the first stage of each amplifier ungrounded. The output of the amplifiers was integrated according to the first method described in (2) before it was put on the two oscilloscopes. The stimuli, as explained, were made to begin at the beginning of an interval of $\frac{1}{10}$ sec. (the interval over which integration took place). The moving film camera was started in each case a few sec. before the beginning of the stimulus and stopped half a min. or so after it. Records from the camera were read in a projection device (2).

This experimental procedure suggests another way of looking at the problem. Telling the *S* to respond at the end of a tone makes the duration of the tone a fore period for that response, a 'filled-in' fore period. A muscular response is already known to occur in a fore period (3). This might or might not have anything to do with the responses to sounds in the non-reaction condition. Therefore the several intensities of tone were used to see whether the stimulus had any more than a simple signaling effect. From this point of view part of the problem is whether stimulus intensity will affect the muscular response of 'set.' In sum, there are two ways of looking at the problem: Will the 'set' condition affect the reflex responses and will different sizes of reflex response affect the 'set' response?

Results

The first step in the treatment of results was to obtain the average response curve for all 24 *Ss* for each of the six principal experimental conditions for both arms. These are presented in Fig. 1. They all have certain features in common, and also resemble a great deal the response curves shown before (1). Some suggestions that come from looking at the curves are:

1. These, like the earlier curves have a peak a few tenths of a second after the beginning of the stimulus. This was named the a-response. The increase leading up to the peak begins at least as

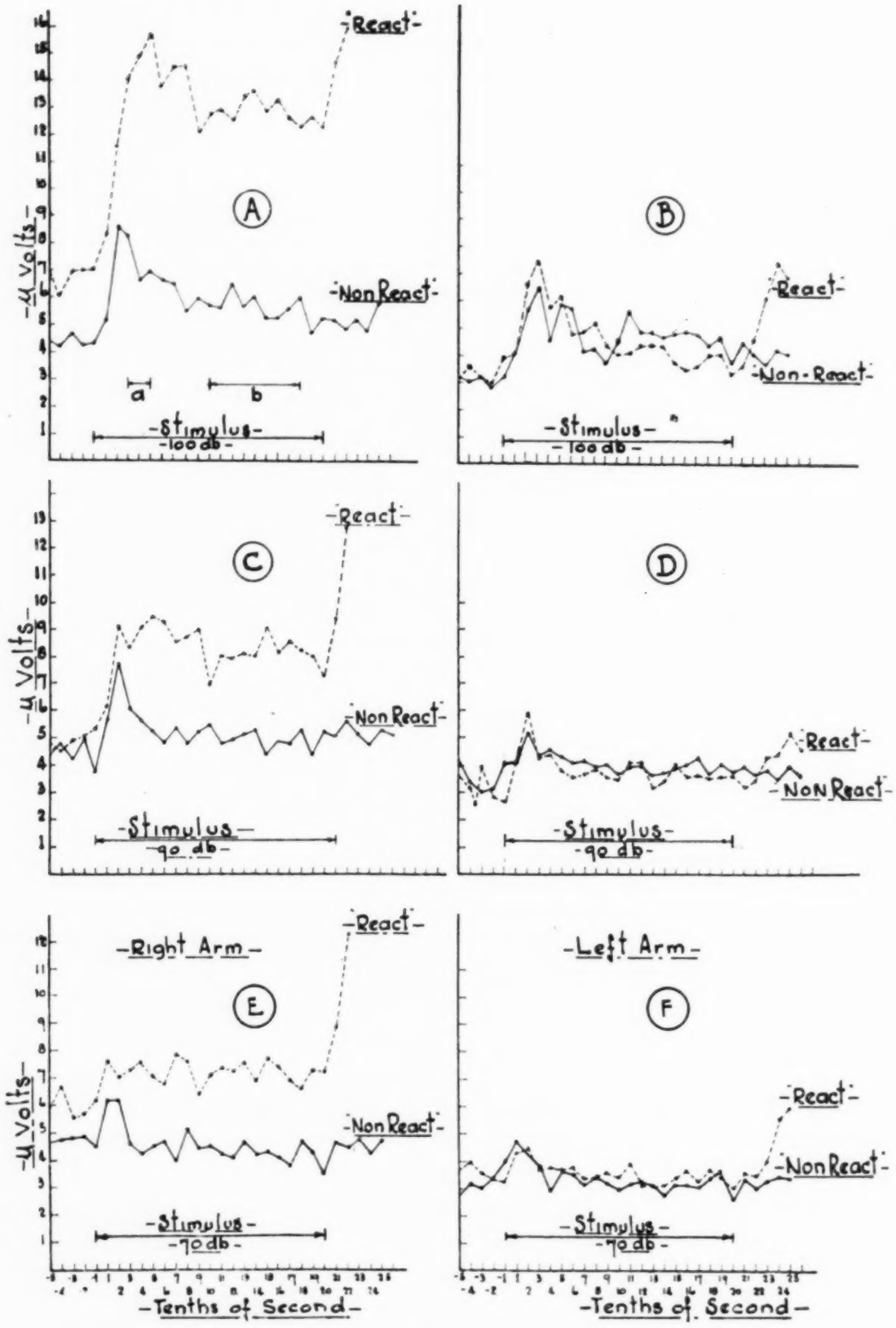


FIG. 1. Mean response curves for different stimulus intensities and instructions

early as the second tenth of a sec. Where the peak is higher, it is reached later; the largest comes half a sec. after the stimulus starts.

2. In some of the curves there is a hint of a second rise after the falling-off of the first. This is the b-response of the earlier study, which, it was thought, would make a hump on the response curve when it is large and when the a-response is not. Generally there is a low point around the end of the first second.

3. All curves for the right are at a higher level than corresponding ones for the left, both before, during, and after stimulation.

In both a and b parts of the response curves the size of the response seems to be related to the stimulus intensity. The relation seems to be there for both left and right and for react and non-react.

5. Both a and b parts of the response curve are raised by the instructions to respond, but only on the right side. If anything, the left is depressed by such instructions.

6. A puzzling difference between these and the curves of the previous study is in size of response and level of tension. Some of the numbers are about the same in the two studies, but they mean different things (**2**). Comparable unintegrated values should be about twice as large as integrated. The readings in this study are considerably larger than those from the former under what seem at first anyway to be comparable conditions.

In the following, these observations will be checked and accounted for so far as possible.

The next step in procedure was to get a measure for pre-stimulus tension and for response. The value in the interval labeled -2 was taken to represent pre-stimulus tension level. It lies between 0.1 and 0.2 sec. before the beginning of the stimulus. For the responses, since the earlier experiment found some different properties in the first and later parts of the response curve, it seemed best to use two representative values—one for the a-response, one for the b-response. To get the first, the greatest of the three readings labeled 2, 3, and 4 was selected for each response, and the corresponding pre-stimulus level (reading -2) was subtracted from it. The value for the b-response was derived in the same way, except that the greatest reading between numbers 10 and 18, inclusive, was chosen. All this is the same as the process used in the earlier study, except that the pre-stimulus measure is moved back 0.1 sec. Then the influence of various conditions upon these measures was checked by analysis of variance. In the tables used for the analysis each entry was the mean of two readings, for there were two responses taken under each possible combination of conditions. The conditions which were systematically varied were:

1. Reaction vs. non-reaction.
2. Stimulus intensity.
3. Position in the series—1st, 2nd, 3rd stimulation, etc.
4. Order of stimuli (the 12 orders already mentioned. For their description, see the following.)
5. Right vs. left hand.

For the reaction data right and left hands were treated separately. For the pre-stimulus figures they were included as divisions of the same table. Hence, there were 144 cells in the one sort of table and 288 in the other. In the study of responses separate analyses were made for the a- and b-readings.

In the pre-stimulus period the mean tension (in interval −2) is 5.3 microvolts for the right arm and 3.1 for the left. The difference between these is significant at the one percent level when compared with the remainder variance. For two reasons this is to be expected. In right handed people, the rest condition of the right arm is one of higher tension. Besides this there was probably some tension produced in the right arm by the position in which it had to be held. The right-left difference does not have much to do with the instructions to respond which were given the *S* half the time, for the right and left arm tensions have about the same ratio to each other when the *S* was told not to react.

Another difference appears between conditions of 'react' and 'non-react' instructions. Respectively, the means are 4.5 and 4.0 microvolts, a difference significant at the five percent level. But this difference is really in the right arm only, where the averages are 6.2 and 4.7. In the left arm there is a tenth of a microvolt in the other direction. This increase in the tension of the reacting arm must come from the instructions given (to keep relaxed but respond at the designated signal) or from the after-effects of earlier responses in the series or from both. The same effect was found in an earlier study (**3**).

A third factor, the one called 'order' is significant at the one percent level. This factor would equally well be called a 'differences between pairs of *S*s,' factor, and that may be what it means here. Each presentation order was used for a different pair of *S*s; so this variable is a combination of two things. As it includes the variance of *S*s, it might be taken as a part of the 'error' variance. If it is, the significance levels for the other variables are unchanged.

The other two variables, position in the series, and stimulus intensity have no significant effect on the pre-stimulus tension. There hardly could be any influence of the latter. The lack of position effect shows that the *S*s' pre-stimulus tension neither increased nor decreased systematically with time.

For the two types of responses, a tabulation of the significance levels found for the four variables in the four analyses of variance is given in Table I.

Where there is a blank, the level is greater than five percent. If the 'order' variance is taken as part of the 'error' the only figure changed is the one marked with an asterisk, which then falls just below one percent. These figures confirm points numbered 4 and 5 in the discussion of the response curves, and show some facts about 'position' and 'order.' These points will be looked into further.

TABLE I

LEVELS OF SIGNIFICANCE FOR THE FOUR EXPERIMENTAL VARIABLES

Variables	a-Response		b-Response	
	Right	Left	Right	Left
Stimulating intensity	1%	1%	1%	5%
Reaction vs. non-reaction	1%		1%	
Position		5%		1%*
Order	5%		1%	1%

The difference between right and left arm responses needs separate testing. From the 'error' variance (including the 'order' factor) the standard errors of the mean a-response and of the mean b-response for each side were computed, and from them the standard error of the differences between sides. These are given in Table II.

Like rest level, the response mean (i.e., increase over rest level) is much greater for the right. It is when the *S* is told to respond that this right-left difference occurs; otherwise the difference is very slight and not significant, though in the same direction.

TABLE II

COMPARISON OF RIGHT AND LEFT ARM RESPONSES

	Right	Left	Difference	*t*
a-response	6.2	3.6	2.6	5.4
b-response	6.0	3.8	2.2	4.4

A striking point about the above averages is the nearly identical size of the a- and b-responses when measured in this way. This is quite contrary to the impression given by the response curves in Fig. 1. The fact is that, taken one by one, the b-responses are as large as the a-responses but do not all come at the same time.[1] Therefore,

[1] However, the b-response reading needs to be discounted by more than the a-response for the inflating effect of chance error.

the curves in Fig. 1 based on average reading at particular times, show, in general, a moderate level maintained for a long time in the b-region. Individual response curves are likely to show a higher level for a briefer time. Those shown in Fig. 2 seem to be representative.

FIG. 2. Response records for single responses. Read right to left. Dotted line signals stimulus duration. Upper record taken under instructions to respond, lower under instructions not to respond. Top line in each record taken from left arm, bottom from right arm.

The effect of stimulus intensity on both a- and b-responses is very clear. Fig. 3 makes a graphic comparison of the trends under reaction and non-reaction conditions. The responses increase with stimulus intensity, apparently in an accelerated fashion. There is no difference between the a- and b-responses in either size or function. This is surprising because of previous results (1). In the earlier study an increase very much like this was found in the intensity function for the a-response, but it was on a lower level, and no such varia-

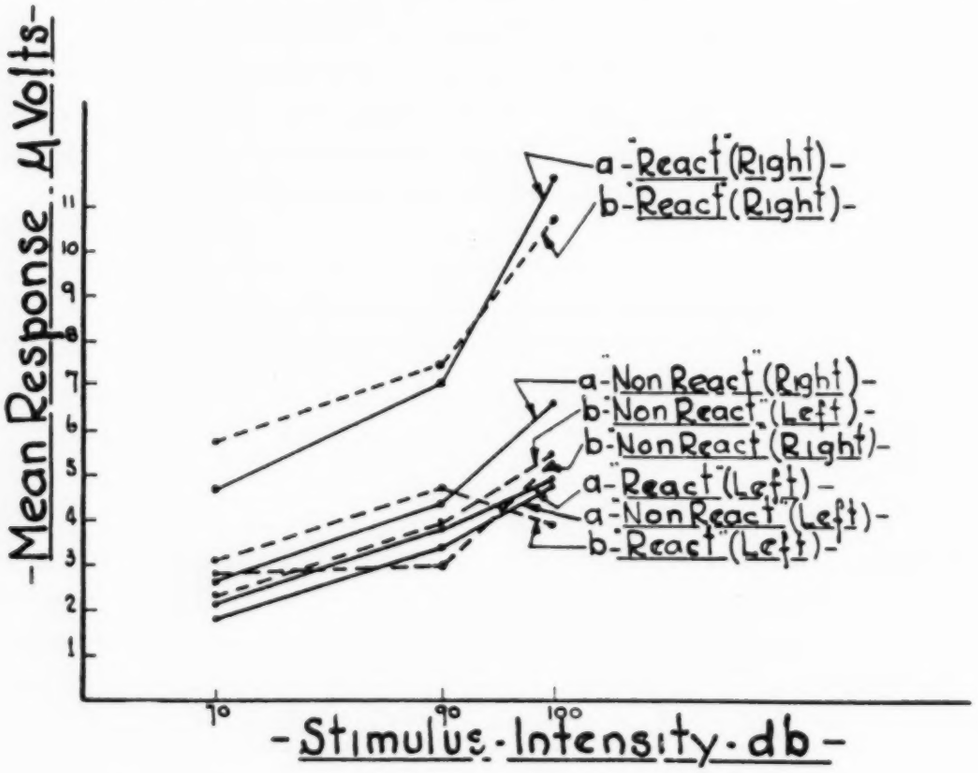


FIG. 3. The a- and b-responses as functions of stimulus intensity

tion was seen in the b-response. This and other divergences of the two sets of results will be taken up below.

In data based on maximum readings, as these are, it is not easy to say when a mean of these is really equivalent to zero, i.e., shows nothing more than a chance deviation. So long as a value can be found which is significantly lower than another the latter must be greater than zero. But is the lowest obtained value really any more than zero? In Fig. 1 the response curves for the weakest stimulus show clear evidence of an a-response. But for the b-response with a weak stimulus under the 'no-response' condition there is no such evidence for either arm. In fact, the mean tension during the b-response period is lower than it is at the point chosen to represent the pre-stimulus state. If there were a significant b-response this could not happen unless there were also a depressant effect to offset it. The data do not show a b-response, therefore, under these conditions.

The effect of instructions to respond is also very clear in Fig. 3. As the analysis of variance brought out, there is an elevation of the fore period effect when the *S*s are told to make a response, but only in the arm which is to make the final movement. Fig. 3 represents the increases in tension during the stimuli. So there is not only an increased tension in the right arm (see the analysis of pre-stimulating period), but there is also a greater responsiveness brought about there by the instructions. The heightening of responses is actually greater for those which are large—those coming from stronger stimuli, that is. The points on the two upper lines can be obtained by adding a constant percentage to the corresponding points on the lower 'non-respond' lines.[2] The influence of stimulus intensity seems to be greater when the *S*s are told to make the final response.

As in Fig. 1 the curves for the left arm in Fig. 3 look as though the instructions to respond might have inhibited the response to the strong stimulus. But a *t*-test shows the difference is not significant.

Both right arm tension (pre-stimulus) and right arm responses are raised by the instructions to react. In the previous study a relation was found between induced tension and response size. If that relation is present here it might explain the increase in the responses. The means to be compared are shown in Table III. Both responses increase in this series as the pre-stimulus level increases with what seems an accelerated rate. When responses to different stimulus intensities are treated separately, it appears that for the high intensity and a- and b-responses are linear functions of pre-stimulus tension. Their deviations from the linear function do not average more than half a microvolt. With the lower stimulus intensities, of course, the relation is less linear.

[2] The constancy of the percentage cannot, however, be proved statistically.

TABLE III

MEAN RESPONSES OF THE TWO ARMS AND THEIR PRE-STIMULUS TENSION

	Pre-Stimulus Level		a-response	b-response
Right arm	React.	6.2	7.8	8.1
	Non-react.	4.7	4.6	3.9
Left arm	React.	3.0	3.6	3.8
	Non-react.	3.1	3.5	3.7

In the former study a pair of equations described quite exactly the relation of responses to pre-stimulus tension directly induced. These took into account the ipsilateral and contralateral tension in the following fashion:

$$a = 1.41 + \frac{9.6}{c} - 5.1$$

$$b = .61 + \frac{7.4}{c} - 1.6.$$

Fig. 4 shows what predictions these equations would yield for the present data. In this diagram perfect prediction would be represented by a line running through the origin at an angle of 45°. The points could very well be regarded as chance deviations from a straight line with a 45° slope, though not passing through the origin.[3] In other words the responses are predicted quite well (with a mean error of less than one microvolt) if the quantity 1.6 microvolts is added to each a-response prediction and 2.0 to each b-response prediction. (The need for these adjustments is not surprising in view of the different methods of recording, and the different stimuli in the two experiments.)

The discrepancies shown in Fig. 4 are small, but they are much larger than remained in the former experiment. There is also a suggestion that they vary systematically. The responses of the present experiment have a relation to initial tension very much like that found in the preceding experiment, but here there may be some additional factors at work. In other words, instructions to react and instructions to be tense have very similar effects on fore-period responses, but perhaps not identical effects.

In the analysis of variance the factor called 'position' would be significant if the responses become greater or less with repetition of the stimuli. For the left hand such a factor is significant. There is just as much variance also in the right arm. Possibly the interpretation should be that the error variance is greater, rather than the position variance is insignificant. But a significant or credible position variance does not necessarily mean adaptation, or even a trend. The means need to be examined in order to see whether they go up or down or fluctuate. There is much fluctuation from stimulus to stimulus, but trends appear when the responses are averaged together in blocks of three. The amount of adaptation is comparatively small, amounting to 25 to 40 percent overall decrease. There is much less

[3] This is not exactly the line which best fits the points in the figure.

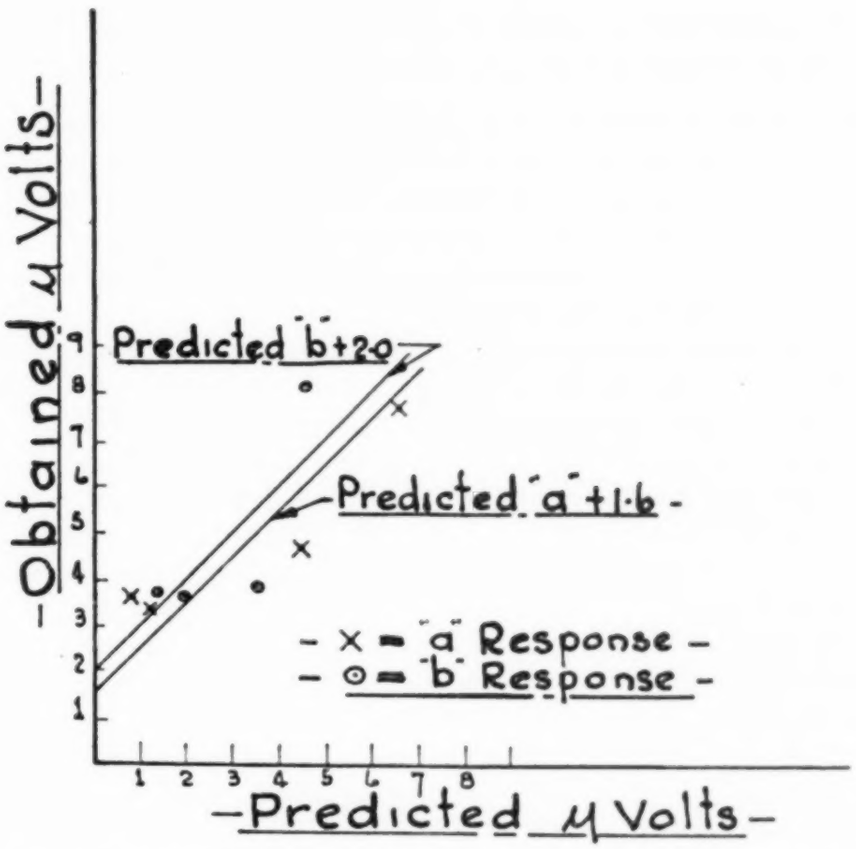


FIG. 4. The obtained data on the a- and b-responses compared with the regression line derived for a previous set of data (1)

than reported for the a-response in the earlier experiments. It is like the trend found then for the b-response. But in that experiment such a trend did not meet significance tests; in this it does part of the time.

DISCUSSION

There is much similarity between these results and those described in the previous study of responses to loud sounds. There are also a number of differences, even when conditions seem about the same.

1. Tension levels and responses ('non-react' instructions) are greater in the present experiment. (In comparing the two, one must allow for the factor of two, which will put these results on approximately the same bases as the previous.) The tension level of the fore period is about 50 percent greater here under 'non-react' conditions than under similar negative instructions before. Responses to the loud stimulus (about the same decibel level as the one in Part 2 in the earlier study) are in this experiment nearly four times as large as those found before.

2. The b-response, as well as the a-response, increases with stimulus intensity. It did not do so in the earlier study.

3. The a-response is much less subject to adaptation than it was in three experimental groups (1, 4).

The first of these differences may explain the other two; when responses are raised to a higher level, they behave differently. In the earlier paper it was remarked that no large adaptation effect exists in the startle pattern to a revolver shot (Landis and Hunt, 5) perhaps because the very large responses have different characteristics. It seems as though the a-response in this experiment is nearing that higher level.

In part, of course, the relation is the other way round, for the a-response, because adaptation makes for a small average response. But even for the first three stimuli the a-responses here are about twice as large as in the former experiment, under apparently similar conditions. In a sense less adaptation and increased size are both signs of greater responsiveness.

The b-response also seems to have different behavior when it is large. As a matter of fact the b-response intensity function shows that large responses are more susceptible to variation with intensity, if the non-linearity of the curve may be taken as a fact. At lower response levels, one would predict from this, variations in stimulus strength would make no difference. (The same should be true of the a-response, if a low enough level is reached.)

If the responses behave differently because of their size, the next question is, what has made them large here? It could be a sampling variation. Although no exact test of sampling effect can be made, the order variable gives some estimate of it here. A rough calculation on the basis of this and a similar variable in the previous experiment makes it seem that the difference in pre-stimulus tension could well be produced by sampling variation while the differences in responses at least approach significance. Aside from the sampling factor there is a feature of the present experimental conditions which might have raised the pre-stimulus mean. Throughout this experiment the *S* was required to keep his right hand raised off the cot and resting on a three in. box which contained the button he was supposed to push under 'respond' instructions. Thus, a tension condition more like that of artificially induced tension may have been brought about. A difference in pre-stimulus tension, however produced, could explain the differences in size and properties of the responses in the two experiments.

So far as this experiment is concerned there is not a great deal of difference in the behavior of the so-called a- and b-responses, and there would not seem much need to distinguish them. But taken together the two experiments do show certain functional differences in them, which warrant treating them separately at least for the present.

In the first part of the paper a parallel was observed between the stimulus duration period here and the fore period of a reaction time experiment. A muscular response is known to occur in the fore period (3). When the *S* is about to react, arm tension in the fore period begins to rise after two or three tenths of a sec. and reaches a peak at some time late in the fore period. Such reactions are therefore coincident with the b-response of this experiment, when the *S* is

under 'respond' instructions. The b-response measurements must represent the same response. Now the b-response, occurring in the same period and behaving in the same fashion with regard to stimulus intensity, is found (for those stimuli) even when the *S* is told not to respond. The effect of telling the *S* to respond, therefore, is not to create a new (preparatory) reaction, but to enhance, in a particular locality, a response which occurs anyway.

The instructions to respond have also the same effect on the a-response, usually thought of as a reflex rather than a voluntary response. With these instructions it is enlarged and concentrated in one region. With respect to instructions, and for that matter the other conditions in the present investigation, there is no difference in the behavior of the so-called reflex and so-called voluntary set.

The manner in which the instructions operate to produce such effects looks quite similar to the way in which directly induced tension operates. Instructions produce a raised level of tension in a particular region before the beginning of the stimulus—that is, before the fore period—and subsequent responses are dependent upon this tension level. Whether the effect of instructions is altogether the same as that of directly produced tension cannot be decided, but it is in the main. Physiologically this means that the prior excitation conditions of the organism will regulate and shape the subsequent responses in a predictable fashion. Where this influence is in the organism is another question, whether in the effector, or in the central sectors supporting the effector action, or in both. The present data can hardly give any information on the point.

The original question now seems to be answered. The setting up of a final response movement by instructions does not merely add a link to an already existing chain. It modifies the preceding links so as to change the character of the whole. This is true at least for the peripheral events, which are probably symptomatic of central process as well. If the individual's judgments are to be understood as showing the effect of stimuli upon him, it must be recognized that the instructions to give judgments are themselves contributing to the effects, which would otherwise be different.

Summary

The problem studied is whether instructions to respond will change the effect a 'sensory' stimulus has upon the individual. The effects studied were the muscular responses in the forearm produced by rather loud tones lasting two seconds. The question is how these are affected by having the *S* make a voluntary finger movement at the end of the sound. Three sound intensities were used and 24 *S*s

were put through various rotations of the conditions. Integrated action potentials were recorded from both forearms. Telling the *Ss* to respond heightened the pre-stimulus tension in the arm to be used. Two stimulus effects, called the a- and b-responses, in the first and later parts of the two-second period were analyzed. The first of these is part of the 'startle pattern,' the second is evidently the 'set' reaction, yet both behave in the same manner with the variable here studied. Both show a slight adaptation. Both vary with stimulus intensity whether the *S* is going to make a finger movement or not. When the *S* is going to execute that movement, these fore-period reactions are very much larger in the arm that is to respond, showing that instructions do modify the antecedent effects. Much, if not all, of this increase seems to take place because the preliminary tension level is higher. Because of this influence of instructions an *S*'s judgment of the stimulus cannot reveal the amount of its effect on him when he is not judging.

(Manuscript received October 27, 1947)

References

1. Davis, R. C. Motor effects of strong auditory stimuli. *J. exp. Psychol.*, 1948, **38**, 257–275.
2. Davis, R. C. An integrator and accessory apparatus for recording action potentials. *Amer. J. Psychol.*, 1948, **61**, 100–104.
3. Davis, R. C. Set and muscular tension. *Indiana Univ. Pub. Sci. Ser.*, 1940, No. 10, 1–30.
4. Davis, R. C. The muscular tension reflex and two of its modifying conditions. *Indiana Univ. Pub. Sci. Ser.*, 1935, No. 3, 1–22.
5. Landis, C. & Hunt, W. A. *The startle pattern*, New York, Farrar & Rinehart, 1939.

A COMPARISON OF TWO INDICES OF PALMAR SWEATING *

BY M. A. WENGER AND J. C. GILCHRIST

Department of Psychology, The University of California, Los Angeles

For many years physiologists and psychologists have been searching for a practicable method for the measurement of palmar sweating. The significance of individual differences in function of this area of so-called 'emotional sweating' has received considerable attention and has been reviewed by Kuno (3), and others. Darrow's perspiration recorder (1) and Miles' micro-recorder (4) are perhaps the most practicable of the direct techniques, but both utilize hygroscopic fibers and require frequent and careful calibration. Darrow (2) showed, however, that the log of the conductance change, as determined by measurements of change in the electrical resistance of palmar skin, bears a direct relationship to amount of sweating from the same area as determined by the perspiration recorder. It is possible, therefore, to make use of log conductance change in an individual as an index of changes in amount of sweat secretion in that individual. The use of differences in palmar conductance between individuals as an index of individual differences in palmar sweating has received no direct validation, but Wenger (7) has shown by the technique of factor analysis that a portion of the variance of a distribution of such data is related to other physiological functions mediated by the autonomic nervous system, and has assumed that this portion of the variance is representative of palmar sweating.

Anyone who makes measurements of palmar conductance on many individuals, however, cannot escape being impressed by the instances in which a relatively high conductance is found for an apparently dry skin, and in which a relatively low conductance is coupled with an apparenty moist skin. He will be forced to conclude, as we have, that palmar conductance is not an entirely satisfactory measure of individual differences in palmar sweating.

In 1944 Silverman and Powell (5) reported a simple colorimetric technique for the measurement of sweat secretion which is relatively objective and which provides a permanent record. They prepared a 25 percent solution of ferric chloride in alcohol and painted it on the skin area to be tested, using a current of air to dry the skin both

* Data for this study were obtained in connection with a research project carried out as part of the AAF Aviation Psychology Program.

before and after application. The part to be tested was then placed in firm contact for a period of three minutes with a prepared paper strip impregnated with a five percent solution of tannic acid. Ordinary mimeograph paper was employed for this purpose; it was immersed in the tannic acid solution for three minutes, then was dried and cut to the desired size.

As sweat is secreted by the applied part, it carries with it the readily soluble ferric chloride which reacts with the tannic acid to form a stain ranging in intensity and hue from a faint grey-blue to a dense blue-black. According to Silverman and Powell the intensity of the stain is directly proportional to the amount of sweat secreted, *but they present no data to support this statement.* They do publish photographs showing four intensities of response which serve as their scale of measurement: 0-*faint*, 1-*moderate*, 2-*strong*, and 3-*intense*.

At the time of their report one of us was engaged in a research involving measurements of individual differences in a number of physiological functions, which included determinations of palmar conductance as an indirect index of palmar sweating. The opportunity was taken to obtain concurrent determinations by the tannic acid-ferric chloride technique. This report provides a comparative study of the two sets of data with respect to (a) test-retest reliability, and (b) validity for measurement of autonomic function.

Subjects and Procedure

The *S*s under study were 251 young adult ambulant male patients of the Convalescent Hospital of the Santa Ana Army Air Base. All had been returned from combat theaters with the diagnosis of 'operational fatigue,' or had been returned under the rotation system and had received the diagnosis of 'operational fatigue' after their return. Most had had combat experience but none was severely ill.

*S*s were scheduled for testing[1] on a volunteer or referral basis, and spent about 45 min. in the laboratory, the first 10 min. of which was allowed for rest and adjustment to the laboratory temperature. During this period the nature of the tests was explained and every attempt was made to put the *S* at ease. Sublingual temperature was then determined, and quantity and pH of saliva were measured. The ferric chloride solution was then painted on the palmar surfaces of the first three fingers of the non-preferred hand, dried, and placed in gentle but firm contact with the tannic acid paper for three min. During measurement the *S* was seated comfortably at a table with the forearms in a resting position, and all pressure of contact was controlled by the *E*. An effort was made to maintain relatively constant pressure from *S* to *S*, even though preliminary work had indicated no perceptible variations in results over a wide range of pressures.

At the termination of this measure, zinc-zinc sulfate electrodes were applied to the palms of both hands and the palmar conductance was determined by a constant current method previously described (8), with the *S*s standing in an upright position in order to control the effect of muscular tension on this measure.[2] The *S*s then reclined on a hospital cot and other tests were made; among them mean blood pressure, mean heart period, and electrical conductance from volar surfaces of the forearms.

---

[1] The complete test battery will be described elsewhere (8).

[2] *S*s were seated instead of standing during the determinations by the stain method, in the belief that muscular tension during the involved passive activity could thus be better controlled.

For some of these *Ss* two determinations of sweating were made on the same day by the stain method, the first as described above, and the second by the same technique but about 30 min. later. Other *Ss* were asked to return later for retesting. The resulting data on test-retest reliability are shown in Table I.

TABLE I

TEST-RETEST RELIABILITY COEFFICIENTS FOR TWO INDICES OF PALMAR SWEATING

Stain Index	*r*	*N*
Retest after 30 min.	.67	84
Retest after 17–64 days	.33	47
Palmar Conductance	***r***	***N***
Retest after 17–64 days	.55	47
Retest after 1 min. (children)	.96	30
Retest after 24 hours (adults)	.70	34
Retest after 90–120 days (adults)	.57	39
Retest after one year (children)	.40	60

RESULTS AND DISCUSSION

Before considering these results it should be pointed out that we have deviated from Silverman and Powell in our scaling technique. It is well known that more than four intervals are desirable in correlational studies if greater discrimination is possible. Colorimeter determinations on a continuous scale would seem to be the method of choice, but suitable apparatus was not available at the time of the study. Preliminary investigations employing several raters did indicate, however, that at least nine intensities of stain might be judged accurately, so nine records of graded intensity were selected from the data for use as standards of comparison. On this nine-point scale, inter-rater reliability in judgment for 76 *Ss* proved to be .86, with no rating of a given *S* differing by more than one point. This nine-point scale was used in all subsequent ratings.

It will be seen in the upper portion of Table I that a given stain index of sweating is a fairly reliable indication of response one-half hour later on the same day, but that individual prediction for periods of 17–64 days (mean = 39 days) could hardly be considered satisfactory for a study of individual differences. To a certain extent this is a reflection of the instability of sweating itself, and is to be seen also in the lower portion of the table which shows results of a somewhat higher order for palmar conductance as determined in several different studies, only one of which has been published (**6**).

The fact that the palmar conductance index appears to be a more reliable measure than the stain index does not, of course, indicate that it is a more valid measure of autonomic function. Information concerning this problem is shown in Table II. Here certain variables have been selected for correlational study which previous factorial investigations (**6**, **7**), have shown to possess some communality of

function. Insofar as they do tend to vary together, the underlying factor effecting that covariance is believed to be the autonomic nervous system. In all of these studies palmar conductance has been found to demonstrate a significant relationship to this factor.

TABLE II

RELATIONSHIP BETWEEN SELECTED VARIABLES AND TWO INDICES OF PALMAR SWEATING

Variable	r Stain Index	r Palmar Conductance
Salivary output	−.01	−.05
Salivary pH	−.16**	−.27**
Heart period	−.01	−.13*
Volar conductance	.13*	.14*
Diastolic blood pressure	.06	.09
Sublingual temperature	.13*	.22**
Palmar conductance	.31**	—

** Significant at the .01 level.
* Significant at the .05 level.
N varies between 248 and 251.

The results in Table II show that, without exception, palmar conductance bears closer relationship to the other autonomic variables than does the ferric chloride-tannic acid stain index. It may be concluded that for the measurement of autonomic functions in terms of these variables, palmar conductance is the more significant of the two indices.

The correlation of .31 between the two indices themselves is of particular interest. In view of the known inadequacies of palmar conductance as a measure of individual differences in sweating, it was not anticipated that it would demonstrate a close relationship to the stain index, but neither was such a lack of correspondence to be expected since both are purported to measure the same function. Although the same area of sweating was not studied here by these two techniques, it is to be doubted that the explanation for their lack of relationship lies in this fact. Sweating from the palmar aspect of the fingers is generally believed to parallel hand-palm sweating. This assumption, however, is a point which may deserve further investigation. A more reasonable conclusion concerning the low degree of relationship between the two indices would seem to be that both are influenced by factors other than sweating.

SUMMARY

A new colorimetric technique for the study of sweating (ferric chloride tannic acid stain) has been compared with the technique of palmar conductance for (a) test-retest reliability, and (b) validity as a measure of autonomic function. An indirect approach was made to the latter problem in terms of the relationship of each index to

other autonomically innervated variables as measured in 251 young adult males.

Palmar conductance was shown to be a more reliable measure than the stain index, and to bear a closer relationship to the other autonomic variables. The coefficient of correlation between the two indices themselves was .31.

(Manuscript received November 10, 1947)

References

1. Darrow, C. W. Quantitative records of cutaneous secretory reactions. *J. gen. Psychol.*, 1934, 11, 445–448.
2. Darrow, C. W. The equation of the galvanic skin reflex curve: I. The dynamics of reaction in relation to excitation-background. *J. gen. Psychol.*, 1937, 16, 285–309.
3. Kuno, Y. *The physiology of human perspiration.* London: J. and A. Churchill, 1934.
4. Miles, W. R. A micro-recorder for measureing skin temperature and sweating in airplane pilots. (Unpublished report)
5. Silverman, J. J., & Powell, V. E. Studies on palmar sweating. Part I, A technique for the study of palmar sweating. *Amer. J. med. Sci.*, 1944, 208, 297–305.
6. Wenger, M. A. A study of physiological factors; the autonomic nervous system and the skeletal musculature. *Human Biol.*, 1942, 14, 69–84.
7. Wenger, M. A. The measurement of individual differences in autonomic balance. *Psychosom. Med.*, 1941, 3, 427–434.
8. Wenger, M. A., & Ellington, M. The measurement of autonomic balance in children: method and normative data. *Psychosom. Med.*, 1943, 5, 241–253.
9. Wenger, M. A. Studies of autonomic balance in Army Air Forces personnel. *Comp. Psychol. Monogr.*, 1948, 19, No. 4, 1–111.

STUDIES IN SPACE ORIENTATION. IV. FURTHER EXPERIMENTS ON PERCEPTION OF THE UPRIGHT WITH DISPLACED VISUAL FIELDS

BY H. A. WITKIN

Brooklyn College

S. E. ASCH

Swarthmore College *

I. Introduction

In previous studies of this series (**1**, **2**, **3**) the role of the visual framework in perception of the upright was investigated. The technique employed in the first two of these studies was to present the *S* with a tilted scene and to have him set a rod within it to the true vertical and horizontal. It was a limitation of this technique that the position of the surrounding scene could not be readily changed, and therefore the variety of frameworks that could be employed with any given *S* was limited.

For the study described here a new technique was developed, which permitted ready changes in the field. The normal visual field was eliminated by working in a completely darkened room, and in its place was introduced a simple luminous frame. Mounted on a shaft at its center, this frame could easily be tilted to either side by any amount. An adjustable luminous rod placed within the frame permitted a precise measure of the *S*'s perception of the upright under different conditions of the visual field. Control data for this rod-in-frame technique are available from a previous dark-room study (**3**), in which the *S*, with the body in various positions, adjusted a luminous rod to the upright in the absence of any visual field.

II. The Experiments

A. Procedure.—The apparatus consisted of a square frame, each side of which was one in. wide and 40 in. long. Within this frame was a rod, one in. wide and 39 in. long. The frame and rod were pivoted on the same center, but mounted on separate shafts, so that they could be rotated independently of each other. A protractor, mounted on the shaft to which the frame was attached, moved with the frame against a stationary pointer, permitting direct readings of the position of the frame. A similar arrangement allowed measurement of the position of the rod. Frame and rod were coated with luminous paint, and, during the experiment, were the only items visible in the completely darkened room. When it was necessary to present the frame alone or the rod alone, a cover, constructed of light plywood, was placed over the item to be eliminated.

---

* These experiments were performed in the psychological laboratories of Brooklyn College.

To take readings at the end of each trial, a small desk lamp was turned on very briefly. During this time the *S* kept his eyes closed.

The *S*'s task in this experiment was to adjust the rod to the true V (vertical) and H (horizontal), with his body and the frame in various relations. Two body positions (erect and 28° left)[1] and two frame positions (28° left and 28° right) were employed, in the four combinations possible. Twelve trials were given under each of these four conditions. On the first half of each set of 12 trials, the rod was initially tilted 28° to the same side as the frame, and on the other half it was tilted 28° to the opposite side from the frame. In the body-erect-frame-left condition, the second set of six trials was omitted, however. Under all conditions, judgments of the V were alternated with judgments of the H, so that an equal number of each was obtained. The order of testing was as follows: (1) body erect, frame left; (2) body erect, frame right; (3) body left, frame left; (4) body left, frame right. Finally, after the conclusion of this series, a control test was given in which the frame was in an upright position and the body tilted 28° left. Six trials, with the rod initially tilted 28° with the body, were given for this condition.[2]

The *S* was brought into the laboratory blindfolded, and placed in an upright position about five feet in front of the rod-in-frame apparatus. His task was outlined to him as follows. He was told that upon opening his eyes he would see a square frame, and within it a rod. His object was to make the rod V or H, as indicated at the beginning of the trial. V was defined as parallel to the walls of the outer building, and H as parallel to the floor he was standing on and the ceiling above him. At the outset of each trial, he was to state whether the rod was in the required position and, if not, which way it was tilted.

When these instructions had been given, the frame and rod were placed in position for the first trial, the lights in the room were turned off, and the *S*, who had kept his eyes closed up to this time, was asked to open them and to report on the position of the rod. If he at once perceived the rod as straight (V or H), he was questioned, to ascertain that he regarded it as straight in relation to the building and not only in relation to the frame. If he perceived it as tilted, he was told that the *E* would move it a little at a time, and that he was to say, 'More,' after each turn until he perceived it as straight, when he was to say, 'Enough.' The *E* then moved the rod in the opposite direction from the reported tilt, about 3° at a time, until the *S* said, 'Enough.' At that point the *S* was questioned, to make sure that he actually regarded it as straight in relation to the outer building. He was then told to close his eyes, the desk lamp was turned on, and a reading taken. The frame and rod were adjusted for the next trial, the lights turned off, and the *S* asked to open his eyes.

After completion of the series with body erect, the *S* with eyes still closed, was helped into a leaning position against a board tilted 28° left. This board had at its base a footrest on which the *S* stood. His left arm was placed around the board on which he was leaning, so that his head rested directly against the board. Although this varied somewhat with the *S*'s height, the head was typically about in line with the center of the frame and rod. The remaining trials were given with the *S* in this position.[3]

Fifty-three *S*s were employed in this experiment. Only 41 of these were given the body-erect and the frame-erect conditions, however.

B. Results.—The method of dealing with the measures obtained for each experimental condition was as follows. First, the amount by which the *S*'s setting of the rod was off the true V and H was determined for each trial. In establishing these values, the direction in which the setting was off—that is, whether it was clockwise or counter clockwise from the true V and H—was not taken into account. The scores for all trials given the *S* under a given condition were

[1] Preliminary tests with body tilted 28° right gave results which were very similar to those for the body tilted left. Accordingly the body-right condition was not given, since to do so would have made the series extremely long for the *S*.

[2] In some cases additional tests were given to follow up particular aspects of performance. These special tests will be described at subsequent points.

[3] For a more complete description of this leaning-board arrangement see (3).

then averaged, the result representing his mean error for that condition. Finally, to obtain the mean error for the group as a whole, the mean errors of the individual *S*s were simply averaged. The value obtained in this way of course represents the mean deviation of the obtained settings of the V and H from the true V and H.

The mean error for each of the body-frame conditions employed are presented in Table I. The distributions of individual scores for these conditions are shown in Fig. 1.

TABLE I

MEAN ERRORS FOR EACH OF THE FRAME-BODY RELATIONS TESTED

Since the result for the two body-erect conditions are very similar, they may be treated together.

Condition	Mean Error
1. *S* Erect	
a. Frame tilted 28° left	6.2°
b. Frame tilted 28° right	5.7°
Combined mean, a and b	6.0°
2. *S* tilted 28° left	
a. Frame tilted 28° left	9.4°
b. Frame tilted 28° right	11.9°
Combined mean, a and b	10.4°
c. Frame upright	4.6°

It may be seen from Table I that, whenever the frame was tilted, errors were made in setting the rod, indicating that the tilted frame influenced the perception of the upright. The effect of the frame was smallest with body erect, where the mean error was 6.0°. With body tilted, the mean error was 9.4° when the body was tilted in the same direction as the frame, and 11.9° when it was tilted in the opposite direction. This difference in results between body erect and body tilted confirms the previous finding (**2**) that an erect position permits the most accurate judgment of the upright and reduces the influence of the visual framework. The larger errors for the frame-tilted-with-body condition, as compared with the frame-tilted-opposite-body condition, reflect a greater influence by the field in the former case. A similar difference between the two conditions was obtained in an earlier study (**2**), and the explanation offered there to account for it seems applicable here as well.

In contrast with these results for the frame-tilted conditions, it was found that when the frame was upright the errors were considerably smaller even though the body was again tilted. The mean error for 42 *S*s was only 4.6°, with the mode at 1°. This mean of 4.6° is lower than the values of 9.4° and 11.9° for the other two body-tilted tests, where the frame was tilted. Actually, the 4.6° figure is very

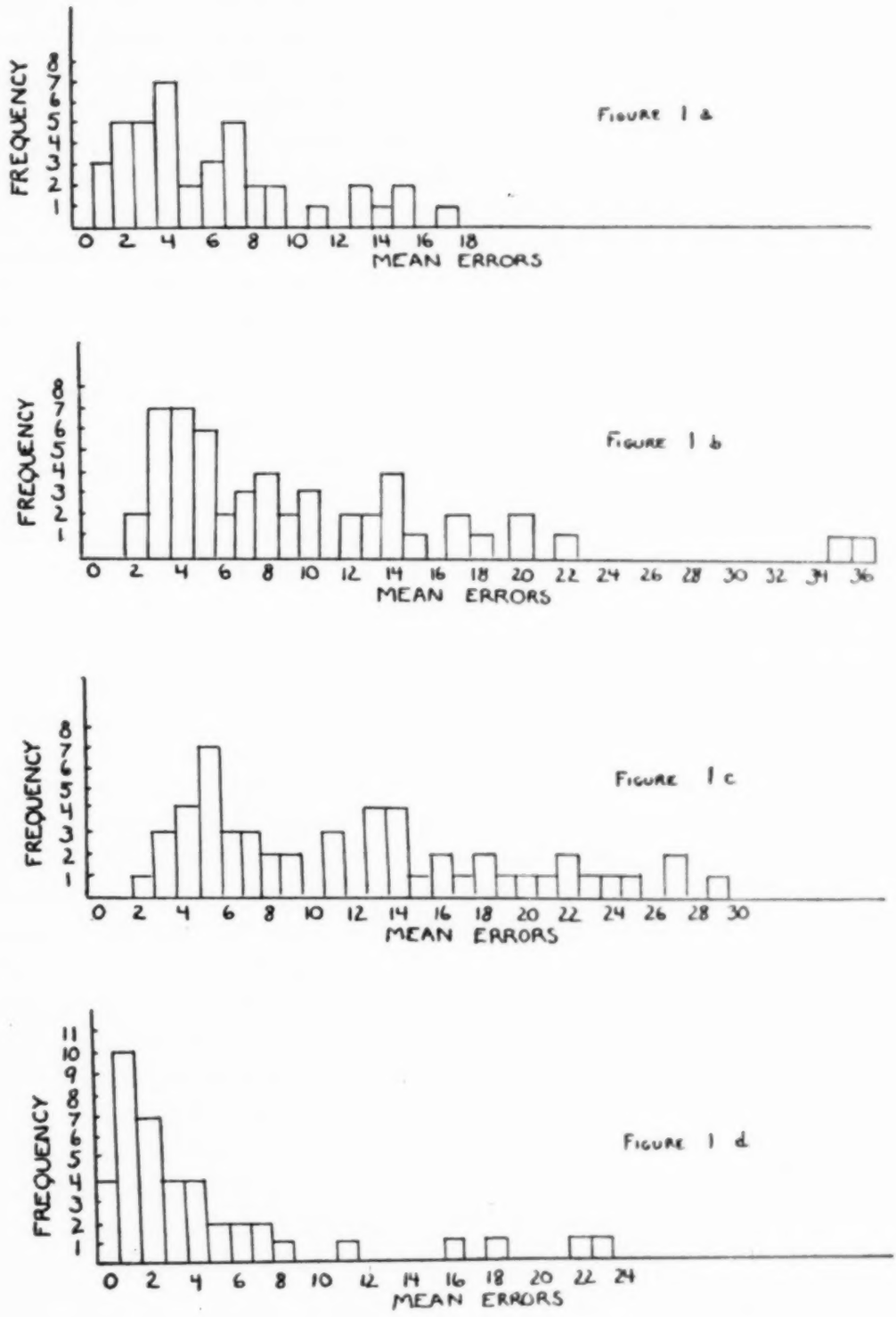


FIG. 1. Distributions of scores

FIG. 1a presents combined results for body erect, frame tilted 28° left and body erect, frame tilted 28° right; FIG. 1b presents results for body 28° left, frame 28° left; FIG. 1c for body 28° left, frame 28° right; and FIG. 1d for body 28° left, frame erect.

much exaggerated by the results of four *Ss* who had extremely large errors. If these four cases are omitted, the mean error for the remaining 37 *Ss* drops to 2.7°. It is clear from these results that the position of the frame greatly influences the perception of the upright.

The effect of the luminous frame is also reflected in the difference between the results obtained here and those obtained in the previous dark-room study (3), where the rod was adjusted in the absence of a field. With body erect, the mean error without a frame was 1.5°, whereas with the tilted frame it was 6.0°. With body tilted, the mean error without a frame was 7.9°; whereas with the tilted frame, it was 9.4° when the frame was tilted with the body and 11.9° when it was tilted to the opposite side. Thus, for each body position, the introduction of a tilted frame caused larger errors in placement of the rod than occurred in the absence of a frame. A consideration of individual performances reveals the same result. For example, with body upright, 33 of 39 *Ss* showed a mean error of 2° or less in adjusting the rod in the absence of a frame. The poorest *S* had a mean error of only 4°. When the tilted frame was present, however, as may be seen from Fig. 1a, most *Ss* had mean errors of considerably more than 4°; and in one case the mean error was 17°.

The influence of the frame is reflected not only in the magnitude of the errors but also in their direction. With body erect, the errors in placing the rod were in the direction of the tilt of the frame in 68.8 percent of all judgments. With body tilted, the errors were in the direction of the frame in 54.5 percent of all judgments when the frame was tilted with the body, and in 80.1 percent of all judgments when it was tilted to the opposite side from the body.[4] Thus the errors in perceiving the upright were most often in the direction of the tilted field.

These findings acquire further significance when compared with some results of the earlier rod-alone experiment. There, with body tilted 28°, the rod was adjusted to the opposite side of the true V or H from the body in 77.6 percent of all judgments. This directional error in the setting of an isolated line, which occurs with small tilts of the body, is the E-effect. In the present study also the body was tilted 28° but the rod was surrounded by the luminous frame. The results clearly indicate that the presence of even so simple a visual

[4] For the body-erect condition, the remaining judgments were distributed as follows. The rod was tilted to the opposite side of the true V and H from the frame in 18.1 percent of the judgments and brought to the true V and H in 13.2 percent of the judgments. For the body-tilted condition, they were distributed as follows. When the frame was tilted with the body, the rod was displaced to the opposite side of the upright from the frame in 38.9 percent of the judgments, and it was brought to the true V and H in 6.4 percent of the judgments. When the frame was tilted to the opposite side from the body, the rod was displaced opposite to the frame in 15.7 percent of the judgments, and in 4.0 percent the upright was correctly established.

field severely weakened and perhaps even eliminated the E-effect. As already noted, with body tilted in the absence of a frame the error was in the opposite direction from the body tilt—in accordance with the E-effect—in 77.6 percent of all judgments; but with the frame present and tilted to the same side as the body, the error was in the opposite direction from body tilt in only 38.9 percent of all judgments. In the majority of adjustments (54.5 percent), the error was in the direction of body tilt—that is, in the direction of the tilt of the frame. When the frame was tilted to the opposite side from the body, the error was in the opposite direction from body tilt, or in the direction of the frame, in 80.1 percent of all adjustments.

It has been seen that the error was more often in the direction of the frame when the frame was tilted opposite to the body (in 80.1 percent of the adjustments) than when it was tilted with the body (in 54.5 percent of the adjustments). A possible explanation of this difference is that when the frame is tilted opposite to the body its influence is in the same direction as the E-effect. Both cause the perceived V or H to shift to the opposite side of the true V or H from the tilted body. Conversely, when the frame is tilted in the same direction as the body its influence is antagonistic to the E-effect. It is reasonable that the error should more often be in the opposite direction from body tilt when both factors work toward that end than when these factors are in conflict. An alternative (or possibly a supplementary) explanation of the less-frequent conformance with the frame when the frame was tilted with the body may be derived from the manner in which the frame was at times perceived under these conditions. Although not investigated systematically, it was established that many *S*s who set the rod in the direction of body tilt and away from frame tilt perceived the frame as an upright diamond, or a diamond tipped slightly toward one side. If the frame is perceived in this fashion and is then used as a basis for judgment, then the rod, to be V must run approximately from the upper to the lower tips of the frame, and to be H must connect the two lateral tips of the frame. When the rod is adjusted in this manner, it is in its final position tilted to the opposite side of the true V or H from the frame. Often, and possibly most of the time, when in the frame-with-body condition the error is in the opposite direction from the tilt of the frame, it would be a mistake to conclude that the frame had not been involved in the judgment. Such an incorrect conclusion results from gauging the main axes of the frame according to their objective loci in space rather than according to the *S*'s perception of them. It is clear that the effect of the frame in the frame-with-body condition is greater than is indicated by the 54.5 percent figure for the with-frame adjustments of the rod.

Fig. 1 shows that there were very marked individual differences in degree of error in perceiving the upright under the influence of the tilted frame. For the body-erect condition (Fig. 1a), the mean errors range from 1° to 17°; for the body-left frame-left condition (Fig. 1b), from 2° to 36°;[5] and for the body-left-frame-right condition (Fig. 1c), from 2° to 29°. It is evident from these ranges that some *Ss* were able to judge the upright fairly accurately, despite the tilted frame, indicating some independence of the field. At the opposite extreme were *Ss* who showed a ready tendency to accept the tilted field as a frame of reference. In fact, *Ss* in the latter category perceived the frame as fully upright on some trials, despite its 28° tilt, and simply aligned the rod with the frame to make it V or H. Such complete adherence to the frame occurred for the most part with body tilted, and was most apt to occur when the body was tilted opposite to the frame.

To evaluate consistency of performance under the different conditions employed, scores for the three basic body-frame relations were intercorrelated, and the following *r*'s obtained: (a) body erect and frame tilted vs. body and frame tilted to the same side, +.53; (b) body erect and frame tilted vs. body and frame tilted to opposite sides, +.50; (c) body and frame tilted to the same side vs. body and frame tilted to opposite sides, +.52; These values are sufficiently high to indicate that a tendency to be influenced by the tilted frame, or to remain independent of it, characterizes a person's perception of the upright under the several conditions employed. The individual differences noted above thus seem to be fairly stable.

In addition to these differences in quantitative results, there were also significant differences in quality of performance. Although most *Ss*, even when influenced by the frame, were able to make judgments without undue trouble, some encountered very serious difficulty. The latter kind of performance, obtained solely with *Ss* who were greatly affected by the frame, was characterized by marked confusion, prolonged effort in reaching a decision, very large errors, and in extreme cases an utter inability to establish the upright. Because of many significant features in their performance, detailed protocols for three such *Ss* are given below.[6]

Subject W: In the series where the body was tilted 28° left, and the frame and rod were also tilted 28° left, the *S* performed as follows. On the first trial he accepted the tilted rod as V at the very outset. In reply to a question, he stated that the rod in that position ran perpendicularly

[5] In the latter case, the rod was tilted to an even greater extent than the frame, which was only 28° from the upright. This resulted from perceptual shifts in the frame, a phenomenon to be discussed later.

[6] The results for two of these *Ss* were not included in the data presented earlier. Because so many special procedures were employed in following up particular aspects of their performance, the testing to which they were submitted deviated too much from the standard series.

from ceiling to floor. On the next trial, when an H adjustment was required, he moved the rod back and forth in a very finicky and anxious manner, and finally aligned it with the H axis of the frame. On being questioned, however, he reported that the frame was tilted. When reminded that the rod was to be made straight with the floor of the room, and not with the frame, he replied that he could not know where the floor was. The trial was repeated; and when the rod was 53° from the true V, in a direction opposite to the tilt of the frame and his own body, he reported that it was V. (His task was actually to establish the H.) In identifying the side of the frame which was nearest the ceiling, he pointed not to the side *a* (see Fig. 2), which was objectively in that position, but to the adjacent side *b*. The positions of the rod in the first V adjustment above, where the frame was perceived in an '*a*-as-top' view, and in the second V adjustment, where the frame was perceived in a '*b*-as-top' view, are illustrated in Fig. 2. With the perceptual shift in the

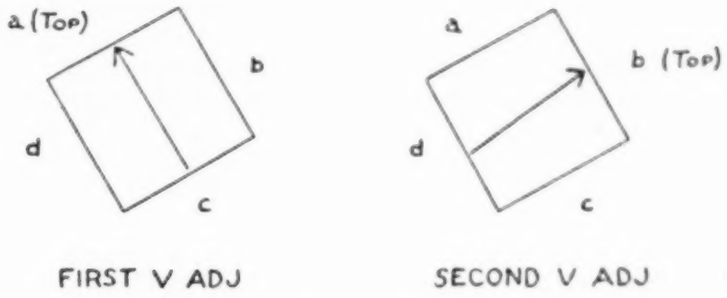


FIG. 2. Adjustments of movable rod by Subject W: (left) when side a was seen as 'top'; (right) when side b was seen as 'top.' (See text.)

main axes of the frame, the perceived upright moved 81°. In the second adjustment, the rod, when perceived as V, was objectively nearer to the H.

On the next three trials, this *S* continued to perceive the frame in its shifted orientation. For the H, he set the rod 55° to the right of the true H; for the V, 55° to the right of the true V; and for the H again, 56° to the right of the true H. The rod was thus successively placed within 7°, 9°, and 6° of the shifted axes of the frame. In each case the *S* asserted that his setting was parallel to the walls or floor of the room.

In another series, the *S* was tilted 28° left, while the frame was tilted 28° left and the rod 28° right. Continuing to see the frame in a shifted position, he misjudged the V and H on four successive trials by 49°, 55°, 52°, and 60°, respectively. The last H setting was 60° to the right of the true H. It was also 2° from the H axis of the frame in its shifted position, and was 15° nearer to the true V than to the true H, which he was attempting to establish. In the series preceding this one, the *S* had made a V adjustment which was 58° to the right of the true H (or 28° to the left of the true V); thus, on these two trials, the *S*s perceived V and H were within two degrees of each other. Under similar conditions with other *S*s, it has been observed that the appearance of the rod would suddenly shift, so that at one moment it looked properly V and at the next moment properly H.

The *S* was then tested with body 28° left and frame 28° right. In the series with the rod tilted 28° right, he accepted the frame as upright on all trials, and therefore also accepted the rod as H and V in its initial position. In the series with the rod tilted 28° left, the frame appeared almost straight, and the rod was brought to within an average of 3° of its axes in the four adjustments made. The *S* thus 'went along with' the frame throughout.

Clearly this *S*'s perception of the V and H was very strongly influenced by the visual framework. Under all conditions he adopted the main directions of the frame completely, or nearly completely. His orientation was also very labile, as demonstrated in the shifts of the visual frame of reference, and this led to large errors and considerable variability in his judgments. For example, he located the V in positions separated by 81°; perceived a line in approximately the same position as V on one trial and H on another; and set the V in positions objectively nearer to the H, and vice versa.

Subject Pe: This *S*, too, demonstrated striking lability of orientation, but less of a glib 'going along with' the frame, and more confusion in his performance.

In one series, the *S* was tilted 28° left, the frame 28° left, and the rod 28° right. On the first trial, the *S* had the rod moved 12° to the right in making it V. In describing the frame, he reported that it appeared like a diamond, but then suddenly stated that it was perfectly straight and had the rod moved to 60° to the right of the true V. In other words, the frame was perceived as upright in a shifted position (i.e., in a *b*-as-top view—see Fig. 2), and the rod was brought to within 2° of its 'new' V axis. On the successive trials, the errors were 51°, 42°, 42°, 44°, and 44°, respectively, to the right of the true V and H. Judgments were made quickly, and with no apparent difficulty.

Errors of the magnitude obtained, as well as the observed shifts in the axes of the frame, can occur only if the position of the body is disregarded. For example, when the rod was accepted as V at a displacement of 60° to the right, it was approximately at right angles to the body, which was tilted 28° to the left of the upright. For the rod to be perceived as upright in the 60° position, the body should have been lying on its side. Reference to the body would have indicated at once that this was not so.

To investigate this problem further, the *S* was asked to make the rod parallel with his body, which was tilted 28° left. This he did very successfully, placing the rod exactly 28° to the left of the upright. He was next asked to make the rod H in relation to the outer building. This time he set it at 40° to the left of the true V, or only 12° from his body, so that his setting was off by 50°. It was pointed out to the *S* that, according to his last adjustment, his body must be only slightly away from the horizontal position. He replied, "Something must be cockeyed," and asked that the rod be moved. Then he suddenly said: "I've lost my perspective of the frame to the room. I can't get my position. I've lost the relation of the frame to my slanting position. I can't straighten it out. I started out by having it."

In his initial judgment of the H, the *S* had proceeded without reference to his body. Then, when he was given the specifically 'postural' task of relating the rod to his body, he did it quite well; but as soon as the task involved more than the body alone, he became confused and unable to proceed. He apparently could not relate what he felt to what he saw. Further, the attempt to establish this relation caused the situation to become very unstable, as evident in his complaint: "I see it one way and then another." It is very likely that 'seeing it one way and then another' reflects a shift from a judgment made in relation to the body to one made without reference to the body.

The *S* was given an opportunity to redetermine the H, and now set it at 50° from the true H, or 12° from his body. He was then asked to raise his body toward the upright. Halfway up, he burst out with a loud exclamation, and began to laugh as he recognized the magnitude of his error. He was returned to the tilted position (28° left), asked to close his eyes and then open them again, and to identify the top and bottom of the frame. He answered: "I can't. I've lost it. I've lost the relation of the frame to the floor." When encouraged to proceed, he did identify the parts of the frame correctly. The test was repeated, and this time the frame was once more perceived in a shifted position, the V now being set at 64° to the right of the true V. The *S* was again asked to raise his body a little toward the upright. He now reported that the frame looked like a diamond. Straightening himself a little more, he reported that it still looked like a diamond. Finally, when he was fully upright, he reported that it no longer appeared as a diamond, but as tilted over toward one side. It is significant that, with the body erect, perception of the frame in a shifted position, whether in a *b*-as-top (Fig. 2) view or a diamond view, cannot be maintained.

Testing was now continued with the frame tilted 28° opposite to the body. In an apparent effort to 'resist' the frame—an attitude prompted by his previous experience—he avoided aligning the rod with it. Instead of moving the rod toward the upright, however, he tilted it even farther than the tilt of the frame. This kind of error reflects the absence of a basis for judgment once the visual frame is excluded from consideration. On four trials, with V and H judgments alternating, his settings were displaced toward the right by 53°, 50°, 65°, and 52°, respectively. Throughout, the *S* proceeded in a hesitant and unconfident manner.

Whatever the basis of these judgments, it persisted even when the frame was excluded. Immediately after the determinations described above, the S was tested with the rod alone. His errors on four successive settings were 42°, 45°, 53°, and 51°, again toward the right. It has been found with other *S*s also that when a given manner of determining the upright has been established with a frame present it tends to persist after removal of the frame. This suggests that a 'remembered' visual field may function as a frame of reference.

In a subsequent interview, the *S* made a number of statements of considerable significance. At one point he remarked: "It's kind of a helpless feeling. Lots of times I couldn't seem to place the floor. The room seemed empty outside the frame. The only way I could do it [make the rod straight] was to get the relation between me and the frame, and sometimes I just couldn't get it."

At another point, he said: "There were some positions where I would suddenly lose the relative position of things. Then I would become upset slightly. If I missed one, it might upset me for a few [trials] afterwards. But it would come back." This comment reflects the emotional involvement often encountered in these tasks.

This *S*'s performance provides another instance of orientation determined primarily by the visual framework and also markedly labile. It further demonstrates the severe disorientation that may result from attempts to modify the predominant visual basis of orientation through the involvement of postural experiences.

Subject P: In a lengthy series of tests, during which she stood erect, this *S* adjusted the rod to the upright very accurately, being only moderately affected by the tilted frame. When her body was tilted, however, she came very markedly under the influence of the frame, and made very large errors. With body 28° left, and frame at the same tilt, she immediately accepted the frame as perfectly upright in the shifted *b*-as-top position (see Fig. 2). On the first V setting, the rod was simply aligned with side *a*, so that it deviated from the true V by 62°. On the next three trials, the frame continued to be perceived in the shifted orientation, though not as fully upright, and the settings were off toward the right by 30°, 55°, and 53°. The most striking feature of these results is the contrast in adequacy of orientation with body upright and with body tilted. The diminished usefulness of postural experiences when the body is tilted, and the increased dependence on the visual field, are most dramatically demonstrated here.

With body again tilted 28° left, the frame was presented to this *S* at a 15° left tilt, and was perceived as straight. The S then reported that she felt dizzy and *did not know in which direction floor, walls, and ceiling were.* She also complained that she *did not know the position of her own body*—whether it was pointing toward the ceiling or toward the floor.[7] There was thus complete disorientation, involving both the body and the world beyond the frame, with accompanying dizziness. At several later points the *S* reported a similar total loss of bearings.

In subsequent tests, the rod was omitted. With body remaining tilted 28° left, and frame 15° left, the *S* perceived side *b* as almost parallel with the ceiling. That is, side *b*, which was at a 75° angle to the ceiling, was seen as nearer to the ceiling than side *a*, which was at only a 15° angle to the ceiling. The frame was then moved clockwise, a few degrees at a time, with the *S*'s eyes closed during movement. She continued to identify side *b* as the top, until the frame reached a position 17° to the right, when she reported that *b* was the right side. The *S* clearly was not aware of this sudden shift in her perception of the frame's orientation. This was demonstrated when, questioned by the *E*, she denied that side *b*, which she now identified as the right side, was the same side that she had on the turn immediately preceding called the top. At several other times when shifts occurred, she did not recognize the changed relation of the sides of the frame, although on some other occasions she did notice it.

When the frame was next presented at a tilt of 6° left, the *S* continued to see side *b* as the top, although this side was now at an 84° angle to the ceiling. The frame was then tilted 15° to the right, and still the *S* continued to perceive *b* as the top. As may be noted from the diagram below, *b* would have had to be moved 105° to be parallel to the ceiling, and was actually farther from being parallel than either side *a* or side *d*, which were 15° and 75° away, respectively. In a later observation it was found that the *S* saw *b* as topmost at an even larger displacement, when it was 112° away from being parallel with the ceiling. (A slight further move of the frame, however, to a point where *b* was at an angle of 116° to the ceiling did cause it to be seen as the right side.)

To summarize, one significant feature of this performance lies in the serious effect produced by tilting the body. Also important were the acceptance of the frame as about upright in a

[7] Another *S* in the group reported the very same experience. He complained a number of times that he had become completely disoriented—that he had lost altogether the ability to determine the angle of anything with relation to the floor. At such times he felt quite unable to make a judgment. At one point he did not know where the floor was, or at what angle his body was tilted. He also stated that he had the impression of the darkness taking on the shape of the square.

variety of different positions, and the shifts in the perceived axes of the frame while it remained in the same objective position. The extreme errors made reveal a striking loss of contact with the main axes of space, and also indicate that the position of the body could not have been taken into account in making these judgments, but that the frame, unstable as it was, constituted the main basis of perception.

Sudden perceptual shifts in the main axes of the frame, such as are described in the above protocols, undoubtedly result from the ambiguous structure of the frame. Its four sides are identical, with no intrinsically defined top, bottom, or side. It is only in relation to some external standard, as assumed proximity to the floor, that a given side becomes top or bottom. This character of the frame permits an interchange of function among its sides.

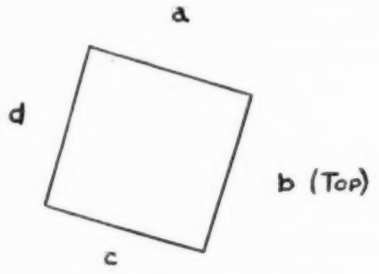


FIG. 3. Position of frame with side b tilted 15° to the right, perceived by Subject P as having side b at the top. (See text.)

The shifts in perception of the luminous frame occurred mainly when the body was tilted. Only rarely were they observed with body erect. An upright position of the body, because it provides a more effective standard of judgment, helps to stabilize the visual framework. Furthermore, these shifts occurred most often when the body was tilted to the same side as the frame. The basis for this is the fact that, though tilted by the same amount, the frame appears somewhat different when tilted with the body than when tilted opposite to it. When tilted with the body, the frame tends to look like a diamond standing on end but tipped over to one side. Though of the two uppermost sides, *a* is clearly nearer to being parallel with the ceiling than *b* (see Fig. 2), the perceptual disparity is not so great as the objective disparity. When tilted in the opposite direction from the body, on the other hand, the frame has the appearance of a tilted square, with *a* unmistakably the side nearest the ceiling. Thus, because perceptually *a* and *b* are more nearly equivalent in the frame-with-body relation, it is easier for an interchange of functions to occur and perceptual shifts are more frequent.

The shifts that occur under these circumstances from the predominant *a*-as-top view are to an upright-diamond view or to a *b*-as-top view. Perception of the frame in the diamond orientation involves, of course, a less radical dislocation of its axes, and seems to

occur as an intermediate step in the shift from the *a*-as-top to the *b*-as-top view. In one case, for example (Subject Pe for whom protocols were presented earlier), the frame was perceived in the *b*-as-top relation. As the *S* gradually straightened his body, his perception of the frame shifted successively to the diamond view and the *a*-as-top view. In another case, the diamond view occurred when, after seeing the frame in the *b*-as-top relation, the *S* was asked to try to see it with *a* as the top.

The fact that when the body is tilted, the visual framework available as a basis of judgment is unstable, may be expected to make for variability in judgments. Because the luminous frame has a very simple structure it provides a rather 'weak' frame of reference. This is additional reason to expect that judgments made in its presence will not be very consistent. To obtain evidence about individual variability of performance, ranges were computed for each *S* for the body-28°-left, frame-28°-left condition.[8] For H judgments, the mean range for all *S*s was 11.3°; for V judgments, it was 12.9°. These results indicate that the rod appeared upright over a considerable range of positions. In fact, *S*s often verbalized this point, complaining that the rod seemed straight in many different positions.

It has been noted that some of the *S*s for whom protocols were presented encountered difficulty in relating the rod to the position of the body when the rod was under the influence of the frame. At times, the frame affected the perception of the position of the body itself, making it even more difficult to judge the position of the rod. Thus it was found in several cases, where the frame became perceptually unstable, that the *S* was unable to establish the position of his own body, being unsure of which way feet and head were pointing. In other instances, movement of the frame produced the illusion that the body itself was being moved. This illusory movement was illustrated most dramatically in a special test where the *S* was in a horizontal body position and viewed the luminous framework as it was being moved slowly through 360 degrees. Under these conditions the *S* reported that the frame appeared upright *in all positions*, despite the fact that he perceived its rotation. At the same time he reported that he felt *himself* turning. As the frame was moved from the upright to 45 degrees, he stated: "I see a perfect square, only I am hanging with my head down; I am the diagonal. As you moved the frame slowly, it seemed as if somebody was picking me up by my feet while my head was sliding down. I saw the frame moving, but all the time it remained horizontal to the floor." This *S* did not experience the induced body movement equally at all points. When the frame had rotated approximately 90 degrees, or, more exactly, to the point where he felt his head hanging down vertically, he experienced suddenly a change

[8] In determining the range, the difference between the *S*'s two most extreme judgments, with account taken of sign, was computed.

of the base line to an altitude, and simultaneously the snapping back of his body to the horizontal position. He reported: "A new line became a base; I felt I couldn't turn my body any more. My body then jerked back to the horizontal position. Up to then it was moving around like the hand of a clock." It is significant that as simple a framework as the luminous frame can exert such an effect upon the perception of body position.

III. Some Further Experiments

A. Judgments of the upright with the frame at different tilts.—In the main experiment, described above, the frame was always tilted 28°. These supplementary experiments employed additional positions of the frame: (a) 15° left, (b) 30° left, (c) 60° left. The *S*s stood erect, and set the rod to the V and H with the frame at each of these angles. Two judgments of the V and two judgments of the H were obtained for each position of the frame from a group of 15 subjects.

The mean errors obtained were as follows: (a) for the 15° tilt of the frame, 8.5° for the H and 8.6° for the V; (b) for the 30° tilt of the frame, 8.9° for the H, and 9.3° for the V; (c) for the 60° tilt of the frame, 5.9° for the H, and 5.9° for the V. In most instances, the errors were in the direction of the tilt of the frame.

Although it did not seem to affect the mean errors in a systematic way, the degree of tilt of the frame did have a striking effect on the range of the errors. Considering results for individual *S*s, the errors are found to range from 2.0° to 15.0° for the frame tilt of 15°; from 2.7° to 26.4° for the tilt of 30°; and from 1.1° to 17.6° for the tilt of 60°. For the first two positions it is seen that the range of individual scores increased with the increased tilt of the frame.

In terms of both errors and ranges, the scores for the 60° position of the frame show a drop rather than the expected increase when compared with scores for the 15° and 30° positions. The basis of this result is suggested by the direction of the errors for each tilt of the frame. With the 15° and 30° positions, errors were always, except for a single *S*, in the direction of the tilt of the frame. With the 60° position, however, some of the errors were in the opposite direction from the frame tilt. If such errors are counted as negative, the range of the errors is from −15.0° to +17.6°. Computed in this way the range is actually larger than for the 15° or 30° positions of the frame. The occurrence of negative errors is explained by the fact that the frame may, in the 60° position, be perceived as either 60° to the left or 30° to the right. Because of the equivalence of the four sides, one mode of perception is just as likely as the other, although the fact that in the two preceding tests the frame had been tilted 15° and 30° left may have caused the same sides to continue their functions as base and altitude, thereby possibly favoring the 60°-left view over the 30°-right view. In any case, depending on which view of the frame is adopted, the direction of the errors may be to one side or the other of the true V and H. The possibility that the frame may serve as a standard in two alternative ways, plus the fact

that some *Ss* shifted from one standard to the other, may reduce the effect of the frame upon perception of the upright in that position. This would help account for the smaller errors found.

B. Judgment of a frame within a frame.—In all the experiments described so far, perception of a very simple visual structure (the rod) within a larger field was studied. It is possible that the appearance of such an item is particularly affected by the surrounding field. In this experiment a more complicated structure—a small square frame—was substituted for the rod, to determine whether perception of it would be more independent of the field. Each side of the small frame was 24 in.

As with the rod, this small frame was pivoted at the same center as the large frame, and the *S*'s task was to make it straight, with the larger frame in a tilted position. Duplicating the conditions of the main experiment reported previously, the body was tilted 28° left and the large frame was on some trials tilted 28° left and on other trials 28° right. The small frame was also initially tilted 28°, either with the large frame or opposite to it. Nine *Ss* were tested. With body tilted left and with the large frame also tilted left, the mean error was 9.0°. This compares with the value of 7.8° for the corresponding rod-in-frame test. With body tilted left and large frame tilted right, the mean error was 15.2°. The value for the corresponding rod-in-frame test was 13.9°. It is clear that perception of the small frame was markedly affected by the larger frame. The errors, though based on a very small number of cases, are in fact slightly larger than those for the rod-in-frame test. Apparently, being within the field of the large frame, perception of the small frame is as much affected as was the perception of the simpler rod.

C. Reproduction of directions with and without a field.—In one variation of this experiment, the *S* viewed the rod alone—tilted 15° to the left—for a one-min. period, with the instruction that he would later have to reproduce its position. The rod was then presented within the luminous frame, which on five successive trials was first straight, then tilted to the left 15°, 30°, 45°, and 60°; each time the *S* was required to place the rod in the initial 15° position. This series was repeated four times, so that four settings of the rod were obtained at each of the five positions of the frame.

In a second variation of this experiment, the *S* had to set the rod at a 45° angle, instead of a 15° angle, with the frame upright, tilted 30°, and tilted 45°. One *S* was used for each experimental condition. These were *Ss* who in other tests involving the luminous frame and rod had proved to be strongly influenced by the position of the frame in judging the rod. While their results here, which were again very extreme, are not representative, they are significant in showing the extent to which some people may depend upon the visual framework.

Following are the results obtained with the *S* for whom the 15° position of the rod was used. When the luminous frame was upright, the reproductions of the 15° tilt of the rod were highly accurate. The four adjustments averaged 12.8°, or only 2.2° off the 15° position required. As soon as the frame was tilted, however, large errors appeared. In every case the rod was set too far in the direction of the tilt of the frame. With the frame

tilted 15°, the reproduced position of the rod had a mean value of 31.0°; with the frame tilted 30°, a mean value of 45.0°; with the frame tilted 45°, a mean value of 56.3°; and with the frame tilted 60°, a mean value of 47.7°. In all positions of the frame except the last one, the rod was placed about 15° farther to the left than the frame. The 20 individual settings made by this *S* in an attempt to reproduce the initial 15° position of the rod varied from 12° to 61°. It is necessary to conclude that this *S* was not capable of reproducing the objective tilt of the rod. What he did reproduce most of the time was an angle that stood in a constant relation to the framework. This *S*, it should be noted, fully understood the instructions and realized that the changing position of the frame was introducing a distortion into his judgments.

Results very similar to these were obtained with the other *S*, who attempted to reproduce a 45° tilt of the rod. With the frame upright the average setting of the rod was 45.6°, the individual settings ranging between 42° and 49°. With the frame at 30° the average setting was 79.5°. Finally, with a 45° tilt of the frame, the average setting was 88.0°, or virtually horizontal. The individual settings for all three positions of the frame varied from 42° to 95°, though the *S* believed throughout that he was tilting the rod at a 45° angle.

These results cannot be ascribed to 'faulty memory. The accuracy of reproduction when the framework is upright conclusively rules out this possibility. The systematic relation between the rod settings and the tilt of the frame warrants the conclusion that the *S*s perceived the direction of the rod principally in relation to the frame.

IV. Discussion

The study described here has provided more dramatic evidence than did our previous ones of the importance of the visual framework in the perception of the upright. This is mainly because the framework employed was a 'weak' one, and therefore perceptually unstable for some people. Since for such people the frame was also typically adhered to as a basis of judgment in each of its shifted orientations, there resulted enormous errors in judgment of the upright. Summarized below are some of the extreme results from performances, described in Section II, of *S*s whose perception of the frame was particularly labile.

In one case a given side of the frame was perceived as topmost when it was objectively nearer to the floor than to the ceiling. In the given position it would have been necessary to move the frame 112° to make it parallel to the ceiling. Another *S* perceived a H rod as perfectly V, so that the axes of space were in effect shifted by 90°. In another case, the rod was judged to be V in positions that were 80° apart. In addition to causing these enormous errors, the perceptual instability of the frame also contributed in some *S*s to marked con-

fusion about their orientation. This was reflected first of all in increasing hesitancy and anxiety in making judgments. It was also indicated by the following kinds of experience, exhibited and reported by *Ss*: (a) inability to judge the relation of rod, body, and frame, to the point of finding it impossible to make any judgment at all; (b) inability to locate the unseen walls and ceiling of the room ("There seems to be nothing outside the frame," or "The darkness outside the frame has taken on the shape of the frame"); and (c) inability to establish the position of the body itself, to the extent of not being sure which way feet and head were pointing. Marked confusion of this kind, and enormous errors in judgment such as those cited above, never occurred in our previous studies, where the framework was more complexly structured and more stable.

Results of this kind were obtained with extreme *Ss*, and were of course not typical of the group as a whole. In fact, in most cases the 'weak' structure of the frame caused smaller errors in judgments of the V and H than were obtained with the frameworks employed in our first two studies. In brief, although the framework used here made possible enormous errors in some extreme cases, it typically did not produce very large errors. This will be evident from a comparison, to be presented later, of results for the present framework and for those used in our first two studies.

As in our first two studies, striking individual differences in performance have been found here. Although perception of the position of the rod was generally affected by the position of the luminous frame, the degree to which it was affected varied considerably among the different *Ss*. To some the tilted frame tended to appear upright most of the time, and the rod was therefore adjusted to its main axes in making it V and H. At the other extreme were *Ss* who were able to make proper allowance for the tilt of the frame and to bring the rod close to the true V and H, suggesting the use of some other basis for judgment. That a given degree of dependence on the visual framework tends to characterize a person's perception of the upright in a fairly general way, is indicated by the correlations of +0.53, +0.50, and +0.52 among the three main body-frame conditions employed.

In several ways the possibility of perceptual shifts in the framework itself made for greater individual differences in performance than were observed in earlier studies. First, in cases where shifts in the frame occurred, and the frame was adhered to in each orientation, there resulted extremely large errors and great variability in judgment of the V and H. Even if the mean error for the group as a whole was not so great with the luminous frame, the range of individual errors was extended by these extreme cases. Second, in this

study there was a distinction between *S*s for whom the framework was unstable and *S*s for whom it was not. This is a dimension of differentation which did not exist with the mirror or room techniques, since with these frameworks perceptual shifts were never observed. Third, the instability of the framework also made for confusion of orientation in some people, of a kind never found with the room or mirror techniques. This is another aspect of performance in terms of which *S*s could be distinguished. To the extent that performances in the luminous-frame situation permit greater differentiation among people, it seems better suited to the study of individual differences than the room or mirror type of situation. A further advantage of the technique is that it requires a much simpler apparatus.

There is a special aspect of the problem of individual differences that merits consideration. Among *S*s whose perception of the position of the rod was strongly affected by the position of the frame, as already noted, there were those for whom the frame underwent perceptual shifts and those for whom it did not. Aside from this difference in the relative stability of the frame, these *S*s also differed in terms of the difficulty they encountered in arriving at their judgments. Some *S*s, even when perceiving the frame in a shifted position, 'went along with it' in a glib, sure fashion. Others, though also greatly influenced by the frame, experienced very marked difficulty. This was seen strikingly in some of the protocols presented in section II. If we seek the basis of this 'troubled' type of performance, the evidence suggests that it occurs when people who are strongly dependent on the visual framework attempt to go counter to appearances and to use the position of the body in making their judgments. The result of such an effort seems to be that the task is changed from a 'natural' one of reporting on direct perceptual experiences to the more 'artificial' one of interpreting these experiences. The latter way of dealing with the problem may lead to even greater errors than the first; for, as noted in some of the cases described, a complete loss of bearings may result.

There are several lines of evidence to support such a view of these 'troubled' performances. One kind of evidence is illustrated by the results with Subject Pe whose protocols were presented in Section II. This *S* at first tended to set the rod toward the axes of the frame, even though the frame was perceived in a shifted position, so that very large errors resulted. The *E* then carried out some special tests, which served to make the *S* more aware of the position of his body and of the untrustworthiness of his visual impressions. As a result, the *S* attempted to take greater account of his position in judging the V and H; but this required very great effort and, in fact, proved unsuccessful to the point of producing a rather complete loss of bearings.

There thus occurred, through efforts to take account of the body, a change from a glib, visually dependent performance to a very troubled performance. If insight into the untrustworthiness of visual impressions may be provided through the efforts of the *E*, as was the case here, it is also possible that the *S* may gain such insight on his own.

A further indication that 'troubled' performances result from difficulty in relating body position to visual impressions is found in the relative ease of making judgments with the frame and without it, as observed in some cases. One *S*, who became seriously confused when the frame was present, gave accurate and rapid judgments when it was removed. Thus the difficulty disappeared when the task was changed from one in which visual and postural experiences had to be related to one in which postural factors alone were involved. The *S* specifically stated that the task was much easier in the absence of the frame. A result of similar import was noted when *S*s who were experiencing great difficulty in judging the V and H were required to align the rod with the tilted body. In this task, which again did not require relating the position of the body to the position of the frame the judgments proved to be much more accurate, and no effort or disturbance was involved.

From our first two studies (**1**, **2**) and the present study, results are available for three kinds of visual framework, graded in terms of structural complexity. The luminous frame used here, consisting of four equivalent sides, was the simplest. The tilted room of the second study provided a more definite framework, containing many verticals and horizontals and a clearly defined top and bottom. Finally, the mirror situation used in the first study presented a full, natural scene, providing the most richly articulated visual field. From the data now available for these three situations, it is possible to determine the effect of 'strength' of the visual framework upon perception of the upright.

Results with body erect and field tilted may be evaluated first. Considering the extent to which judgments of the V and H deviated from the true V and H, the following mean errors are obtained: for the luminous frame, 6.0°; for the room, 14.9°; and for the mirror, 22.0°.[9,10] There is thus a progressive increase in the amount of displacement of the rod. Equally striking results were obtained when

[9] The results for the room are taken from Situation a (**2**), and the results for the mirror from the with-tube condition (**1**).

[10] The degree of tilt of the framework was not the same for the three situations. The luminous frame was tilted 28°, the room 22°, and the mirror 30°. That differences in tilt of the frameworks did not account for the differences in magnitude of the error is clearly indicated by the finding that the errors were larger with the room than with the frame, though the tilt of the frame was greater.

the body was tilted as well as the frame. Data are available for the luminous frame and the room only, since the body-tilted condition was not used with the mirror. In the case of the room, the mean error was 19.4° when room and body were tilted to the same side, and 20.1° when they were tilted to opposite sides. In the case of the luminous frame, the corresponding values were 9.4° and 11.9°. It should be noted that in the luminous-frame test the frame and body were both tilted 28°, whereas in the room test room and body were tilted only 22°. Typically, larger body tilts lead to greater acceptance of the visual frame work, and larger tilts of the field lead to greater displacement of the perceived upright.[11] On the basis of degree of body tilt and field tilt alone, the errors should have been larger with the frame than with the room. The fact that the errors were considerably larger with the room situation therefore demonstrates in an impressive way that the difference between the two sets of results was determined by the more structured character of the room. The results obtained both with body erect and with body tilted indicate that the more highly structured the surrounding visual framework the greater its influence on perception of the upright.

Greater strength of the framework not only caused the perceived upright to be displaced farther in its direction, but also resulted in greater consistency among successive judgments. This is clearly indicated in the mean ranges obtained with these different frameworks under the same body-framework relations. With body erect and framework tilted, we find for the mirror situation a mean range[12] of 4.6° for V judgments;[13] for the room situation, a mean range of 2.0° for H judgments and 2.6° for V judgments; and for the luminous-frame situation, a mean range of 8.8° for H judgments and 9.2° for V judgments. The variability with the room and mirror is quite small though somewhat larger with the mirror. The variability with the luminous frame is very much greater than with either of the other two situations. With body tilted and framework tilted to the same side, similar results are found: for the room situation, mean ranges of 2.0° for H judgments and 2.6° for V judgments; for the luminous-frame situation, mean ranges of 11.3° for H judgments and 12.9° for V judgments.[14] Again, the more highly structured framework pro-

[11] Evidence for these generalizations will be presented in a later paper.

[12] The mean range was computed by determining the range for each *S*, and then taking the average of these. In the scores from which the *S*'s range was determined, the direction of the error was taken into account (that is, whether it was clockwise or counter clockwise from the true V and H).

[13] Data for H judgments are not available, since the *S* set the rod to the V only in that experiment.

[14] The mirror framework was not tested with body tilted.

vided by the room leads to markedly greater consistency in judgments. The same result is indicated by the direction of the errors for each situation. With body and framework both tilted to the same side, errors in judging the V and H were, in the room situation, always in the direction of the tilt of the framework. In the luminous-frame situation, on the other hand, though errors in placing the rod were most frequently in the direction of the framework, they often occurred in the opposite direction also. The latter kind of adjustment, as noted earlier, depends upon perceiving the frame with its axes shifted. Thus with the luminous frame the rod not only appears straight over a considerable range of positions but it may appear straight to either side of the true V or H. With the room, on the other hand, the rod appears straight within a very narrow range of positions and only when displaced to that side of the true V and H to which the axes of the room are displaced. It is clear from all these results that when the surrounding framework is more highly structured, perception of the upright tends to be more stable.

V. Summary

The effects of visual frameworks of different tilt and of different bodily positions upon perception of the upright were investigated in 53 adult *S*s. The visual field consisted of a simple luminous frame contained in a completely darkened room. Within the frame was a luminous rod which the *S* had to set to the vertical and horizontal. The frame was tilted 28° right, 28° left or was erect; and the body was either erect or tilted 28° left. Tilting of the frame caused a shift in the perceived upright in the direction of the frame. The influence of the frame was smaller with body upright than with body tilted. As in previously reported investigations, individual differences were found. There were *S*s who, despite the tilt of the frame, brought the rod close to the true vertical and horizontal; at the other extreme *S*s perceived the tilted frame as upright, and aligned the rod with it. That these modes of perceiving the upright are characteristic of the individual is indicated by the substantial correlations obtained among performances under the different conditions. The 'weak' structure of the visual framework used resulted in smaller errors, as compared with previous studies, when the rod had to be set to the upright under the influence of the tilted frame. At the same time, the framework was relatively unstable, the top coming to be perceived as a side, and so on. Because of this in some individual cases errors very much larger than those found in previous studies were obtained. Thus, one *S* perceived the horizontal as vertical; another placed the vertical on successive tests in positions 80° apart.

In some instances, the *S* lost the ability to locate the unseen surrounding room, to establish the position of the body, or to bring the rod to the true upright. Differences were also observed in the ease with which *S*s arrived at their judgments. A comparison of the results of the present and preceding study indicates that the effect of the visual field upon the perceived upright tends to be stronger and more consistent, the more richly articulated the field.

(Manuscript received for immediate publication July 27, 1948)

References

1. Asch, S. E., & Witkin, H. A. Studies in space orientation: I. Perception of the upright with displaced visual field. *J. exp. Psychol.*, 1948, **38**, 325–337.
2. Asch, S. E., & Witkin, H. A. Studies in space orientation: II. Perception of the upright with displaced visual fields and with body tilted. *J. exp. Psychol.*, 1948, **38**, 455–477.
3. Witkin, H. A., & Asch, S. E. Studies in space orientation: III. Perception of the upright in the absence of a visual field. *J. exp. Psychol.*, 1948, **38**, 603–614.

STUDIES OF TRACKING BEHAVIOR. II. THE ACCELERATION PATTERN OF QUICK MANUAL CORRECTIVE RESPONSES *

BY FRANKLIN V. TAYLOR AND HENRY P. BIRMINGHAM

Psychology Section, Radio Division III, Naval Research Laboratory

Introduction

Previous studies have shown that with quick arm motions, rate of movement increases with the size of the excursion. Searle and Taylor (7) have shown that this is true of corrective motions made when tracking a target which is suddenly displaced. Barnes (1) reports the velocity of hand motions made in assembling links for a portable typewriter increases with the extent of hand motion. De Montpellier (4) obtained a similar result when studying key tapping movements and translation motions of the hands. Peters and Wenborne (6) report the same finding for movements made with a stylus while Slater-Hammel and Brown (8) observed an increase of rate with distance when moving a sliding pointer from one mark to another.

The present study is a continuation of earlier work and represents an attempt to obtain additional information regarding the manner in which quick, corrective tracking motions are executed. This information has been sought as a step toward building a factual basis for an understanding of the psychological problems involved in target tracking.

The study was designed to analyze quick corrective motions of different extents in terms of rate, acceleration and the third derivative of position. A nearly frictionless tracking control was employed so that the acceleration and third derivative information would be directly interpretable in terms of patterns of force applied by the subjects to the control.

Apparatus

The apparatus is shown in a schematic diagram in Fig. 1. It consists of two parts, the presentation unit in one room and the recording unit in another.

A. Presentation unit.—The presentation apparatus consists of (1) a cathode ray tube upon which the target appears as a green dot against a vertical hairline, (2) a 6-diopter collimating lens

* The opinions or assertions contained herein are the private ones of the writers and are not to be construed as official or reflecting the views of the Navy Department or the naval service at large.

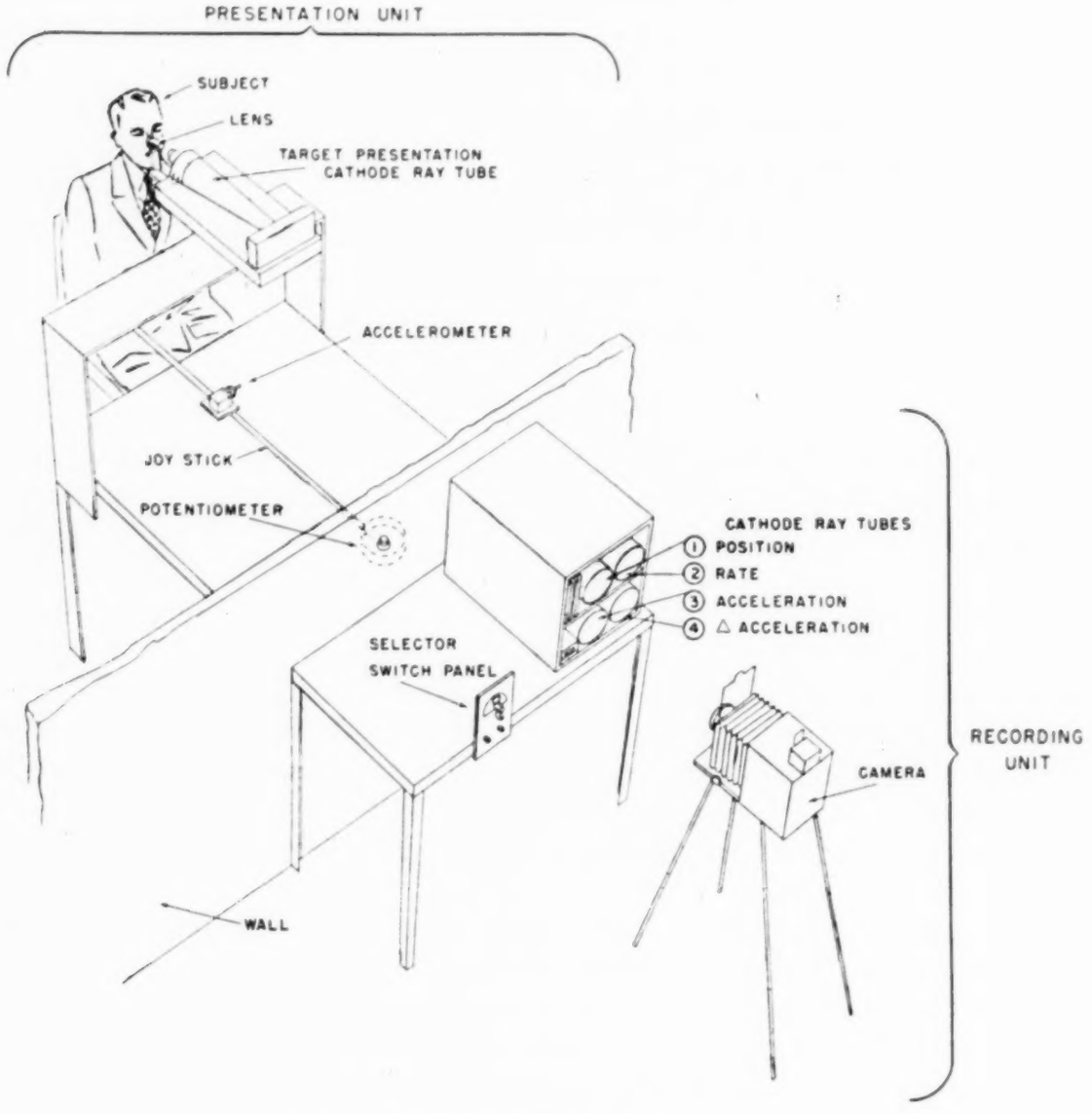


FIG. 1. Schematic drawing of the apparatus

mounted in front of the tube face in such a way that when looking through it, the *S* sees the target and hairline at infinity, and (3) a joy stick which positions a potentiometer and which carries an accelerometer.

The joy stick is 67 in. long and is so mounted that it may be moved horizontally with no more than 6 grams of sliding friction. The output of the potentiometer which is positioned by the joy stick is used for three different purposes: (a) It is fed to the horizontal deflecting plates of the presentation cathode ray tube to control the position of the target. (b) It is fed into the recording unit and mixed with other signals to give a measure of both target and joy stick position. (c) It is electrically differentiated in the recording unit to give a measure of rate of joy stick motion.

The accelerometer is mounted on the joy stick 17¼ in. from the tip of the handle. The accelerometer consists of a 16-ounce Statham strain gauge with a weight mounted rigidly to the armature pin. This weight, plus the weight of the armature and pin, equals exactly one ounce. The strain gauge is mounted on the joy stick with the pin horizontal and with the weight being entirely supported by the pin. With this arrangement, any horizontal acceleration of the joy stick causes the weight to exert a force upon the pin which is proportional to the acceleration. The electrical output of the strain gauge is fed into the recording unit where it is used as a direct measure of acceleration and where it is electrically differentiated into the third derivative of position (Δ acceleration).

The *S* sits on a stool before this apparatus viewing the target with his right eye and with his left eye open in a blanked out eyeguard. As he watches the target he sees it jump away from the hairline to the right or to the left. He shifts the target back to the reference line by moving the nearly frictionless joy stick in the appropriate direction. (A motion of the joy stick to the right causes the target to move to the right.)

B. Recording unit.—The recording unit contains all of the electronic apparatus required for power and signal amplification and for the differentiation of position and acceleration. In addition, it contains a device for causing the target in the presentation unit to shift instantaneously to the right or to the left. A six-position selector switch allows the *E* to select the direction and size of the target displacement. Three sizes of displacement increasing in geometric order may be selected for either direction. These are 1.5°, 3° and 6° visual angle measured at the presentation unit and they require a lateral motion of the joy stick handle of 2¼, 4½ and 9 in. respectively. These target shifts are referred to in this article as numbered 1, 2 and 3 in order of increasing size and are designated L or R depending upon the direction of the displacement as seen by the subject.

Four cathode ray tubes serve as the media by which the electrical signals modified by the *S*'s response are transformed into meaningful pictorial representations of position, rate, acceleration and Δ acceleration. The electron beam in each of these tubes is caused to sweep from left to right across the face of the tube at a rate of 12.5 cm./sec. The sweep is started and stopped in all four tubes simultaneously. The beams start sweeping at the left side of tube 100 msec. before a target displacement is presented to the *S* and the sweep is terminated at the right side of the tube face 1.0 sec. after it starts.

Different signals are fed to the vertical deflecting plates of each of the four tubes. Tube 1 receives the signals which represent joy stick position; tube 2 receives joy stick rate; tube 3, joy stick acceleration; and tube 4, the third derivative of position. Thus, each time the *S* is presented with a target displacement, his response to this displacement is transformed into separate time plots of position and the first, second and third derivatives of position.

A single frame photograph of all four tube faces is taken on 35 mm. film each time the *S* makes a corrective motion. The photographs are taken by opening the shutter before the response starts and holding it open until the trace has traveled across the tube face. The measures made from these photographs provide the raw data for this article.

Experimental Procedure

Ten adults (six men and four women) served as *S*s for the experiment. Three of the men and two of the women had had approximately one hour of previous practice on the apparatus; the other three men and two women had received no practice on the device.

Each *S* was presented with 60 positional shifts of the target, 30 to the right and 30 to the left. Ten displacements of each of the three magnitudes were included in the rights and the lefts. The displacements, randomized as to direction and size, were presented to the *S*s in two sessions. Each session began with an unrecorded 12-trial practice series composed of two shifts of each size to the right and two of each size to the left arranged in random order.

A warning bell was sounded from one-half to two seconds before each target displacement in order to alert the *S* and to prepare him for a response. The time between the warning and the target jump was varied to prevent the *S* from reponding directly to the warning signal.

At the beginning of the first session the *S*s were read the following instructions which were aimed at inducing a set for speed rather than for accuracy:

> This is an experiment on the rapidity of hand movement. You are to watch the green dot on the cathode ray tube while holding the joy stick in your right hand. A short time after the bell rings the dot will jump to the right or left. You are to try to return the dot to the hairline as quickly as possible by moving the joy stick in the appropriate direction. Before the bell rings for the next trial the dot will jump back to its original position. When this happens, move the joy stick to bring the dot onto the hairline. This return motion may be made slowly since it will not be recorded.

When it had been determined that the *S* understood his task, he was instructed to hold the joy stick with the right thumb on the side of the joy stick nearest the body and with the other fingers of the hand held in a natural and comfortable position. The practice series was then started as soon as the subject was ready.

Results

A. Typical response patterns

Though differences between individuals and within individuals appeared in the records, the response pattern of different Ss elicited by target excursions of the same size and direction were sufficiently similar to allow the construction of generalized position, rate, acceleration and Δ acceleration time plots. Accordingly, in Fig. 2 are shown smoothed representations of the typical cathode ray tube time plots obtained for the largest sized excursions (R_3). Except for the letters and the dotted lines representing base lines, diagrams 1–4 are idealized pictures of what actually appears and is photographed on the faces of the four tubes when a subject makes a response to an R_3 target excursion. The differences between these response patterns and those obtained with shorter excursions will be described later.

In viewing these diagrams it should be borne in mind that all four plots have the same time base. To simplify reading from one plot to the next, the same capital letters in each of the four diagrams are used to represent the same points on the time base. The lower case letters represent different times on different plots.

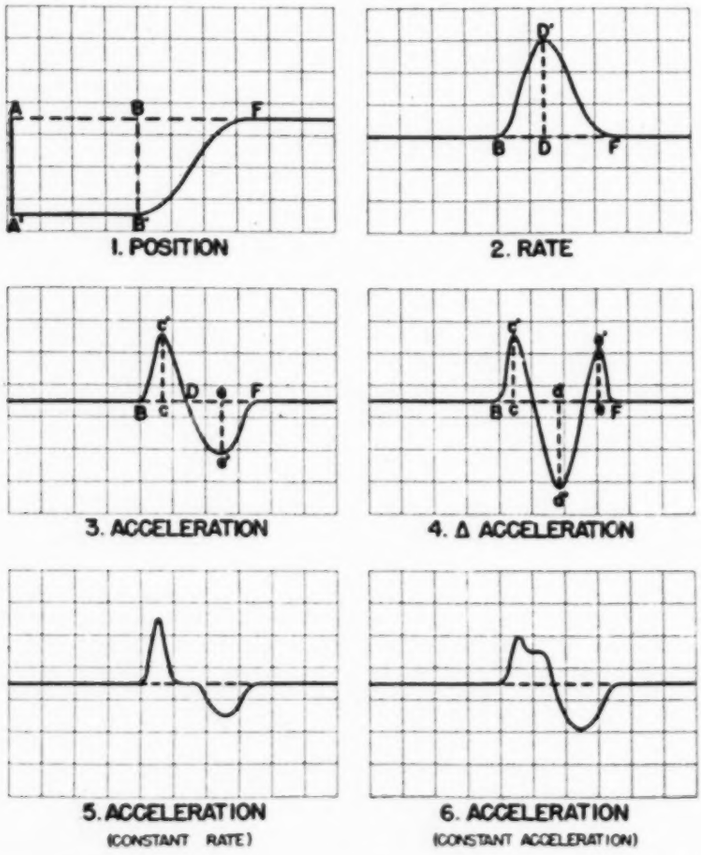


FIG. 2. Diagrams 1–4: Smoothed representations of the typical cathode ray tube patterns produced when the subject responds to an R_3 target displacement

Diagrams 5–6: Atypical patterns found in less than one percent of the data

The interpretation of these time plots is as follows:

(1) *Position.*—(Diagram 1, Fig. 2) The target is displaced at A. The horizontal distance between A′ and B′ is the *S*'s reaction time (0.435 sec.). The response to the target displacement takes place from B to F (0.352 sec.). No overshoot or undershoot is shown in the diagram though both over- and undershoots are frequently encountered.

The value of the reaction time found in this experiment is considerably higher than that observed in previous studies (**6**, **7**). The explanation for this is not altogether clear though it is known that the high value reported here is not a result of faulty measurement. It may be that the tracking arrangement employed in this experiment added an element of complexity which increased the reaction time. In other studies, tracking has been done by moving a pointer to coincide with the position of the target. In the present experiment the target was displaced horizontally by the preset error signal and the error was removed by moving the target back to a stationary hairline by means of the joy stick. This is in a sense an 'unnatural' type of tracking arrangement though it is typical of military instruments. Tests are projected to determine if this is the cause of the long reaction time.

(2) *Rate.*—(Diagram 2, Fig. 2) On the rate plot, the response starts at B. Maximum rate is represented as the vertical distance between points between D′ and D. The response is terminated at F.

(3) *Acceleration.*—(Diagram 3, Fig. 2) The horizontal distance c′c represents maximum positive acceleration and ee′ represents maximum negative acceleration. Point D is the point of maximum rate.

(4) Δ *Acceleration.*—(Diagram 4, Fig. 2) The horizontal distance c′c is a measure of maximum rate of change of increasing positive acceleration; dd′ represents maximum rate of change of acceleration between the points of maximum positive and maximum negative acceleration; while e′e represents maximum rate of change of decreasing negative acceleration.

In describing the course of acceleration shown in Diagram 3 it may be said that the subject accelerates first at an increasing rate (from B to c′), then he continues to accelerate but at a decreasing rate (from c′ to D). Then, instantaneously, he begins to decelerate at an increasing rate (from D to e′), and finally, he continues to slow down but at a diminishing rate until he stops (e′ to F).

Two facts should be noted concerning this typical acceleration picture. In the first place, the accelerating and decelerating portions of the response are not symmetrical. Secondly, acceleration varies continuously; there being no periods of constant rate.

In regard to the differences between the accelerating and decelerating phases of the response, it should be noted that the portion of the curve in Diagram 3 which rises above the base line reaches a higher maximum amplitude than that portion of the plot below the base line, i.e., cc′ is greater than ee′. Furthermore, the upper portion of the curve is subtended by a shorter section of the time base than is the lower part of the curve (the distance from B to D is shorter than that from D to F). The actual times involved are on the average .143 sec. for BD and .209 sec. for DF. Since the part of the pattern above the base line represents positive acceleration while the part below indicates deceleration or negative acceleration, the results indicate that the typical subject decelerates more slowly than he accelerates.

This finding is contrary to that of de Montpellier (4) who observed that the accelerating phase of the movement was typically longer than the decelerating phase. The discrepancy between our findings and his are very possibly the result of the fact that he was studying arm motions of a type different from ours. The movements analyzed in his study were: (1) simple translative motions of the hand involved in moving a finger from one typewriter-type key to another some distance away, (2) hand movements involved in moving over triangular and quadrangular courses and (3) movements of translation of objects of various kinds such as a block of wood, a glass containing water, an empty glass, a block of lead.

In contrast to these motions, the movements analyzed in the present study consisted of horizontal corrective motions which were intended by the *S*s to bring a target to rest upon a hairline. Thus, in our study, though speed was stressed in the instructions, a high accuracy of stopping was also called for. Equivalent accuracy was probably not required in de Montpellier's study and the difference between our results and his may be explicable on this basis. That this is the probable explanation is indicated by de Montpellier's observation that though in general the accelerating phase of movement took longer than the decelerating phase, in the case of translation of a glass of water the asymmetry was reversed with the deceleration taking longer than acceleration. The act of stopping a moving glass of water probably induced in the *S*s a caution which acted somewhat similarly to our own *S*s' 'set' to arrest the target as close as possible to the hairline.

The second noteworthy fact concerning the typical acceleration pattern is that it shows no periods of constant or relatively constant rate or acceleration. In the very few cases where a constant rate was recorded the acceleration pattern appeared as shown in Diagram 5, Fig. 2. Here, the period of constant rate is represented as a hori-

zontal straight line superimposed upon the base line. Upon the few occasions where a constant acceleration occurred, it was characterized by a horizontal straight line parallel to but not on the base line. A representation of this appears in Diagram 6, Fig. 2. Both of these acceleration patterns are rarely found, being observed in less than one percent of all records. *Relatively* constant rates and accelerations were likewise atypical and were found in only a few cases.

The explanation of these atypical findings is not entirely clear though it is quite possible that the occasional constant rates or accelerations obtained were the result of wrist or finger movements partially cancelling out the acceleration of the forearm.

B. Response to target excursions of different size

Nine different measures were taken from the photographic records. Though responses to the right differed somewhat from responses to the left, the *changes* in response pattern with changes in the size of excursion were alike for both directions. Consequently, it is necessary to present only the data for excursions in one direction to illustrate all the major findings of the study. For this reason, only the data for R responses will be reported.

Fig. 3 shows eight graphs representing the changes in eight of the nine measures with increasing size of excursion. The part of each graph title in parenthesis makes it possible to refer to the sample time plots in Fig. 2 and to determine the precise measures upon which these graphs are based. Thus, 2 D′D in the title of the maximum rate graph signifies that maximum rate was measured as the distance D′D as shown in the second diagram in Fig. 2; 3 c′c refers to the distance c′c in the third diagram of Fig. 2. The two time measures could have been made on any one of the four time plots, though actually they were obtained from the acceleration plot (Diagram 3, Fig. 2).

With the exception of the three graphs of maximum Δ acceleration, the graph titles are self-explanatory. The Δ acceleration plot (4 c′c) represents maximum rate of change of increasing positive acceleration; the second Δ acceleration plot (4 dd′) shows maximum rate of change of acceleration between the points of maximum positive and maximum negative acceleration; and the third (4 e′e) shows maximum rate of change of decreasing negative acceleration. In all of the graphs R_1 on the abcissa represents the smallest target displacement, R_2 the next sized displacement, and R_3 the largest.

It is apparent from these graphs that all measures tend upward as the excursions increase in length. Thus R_2 in all cases is associated with a higher value than R_1, and R_3 gives a higher value than R_2. This means that rate, acceleration, rate of change of acceleration and time all increase with increased arm excursion. Furthermore, it is

apparent that the increases are nonlinear since every curve shows a downward bend.

Table I shows the *t*-ratios which express the significance of the difference between neighboring points on the eight graphs. It is clear from this table that, except for the time measures, all differences are significant at the three percent level or better. In the case of the

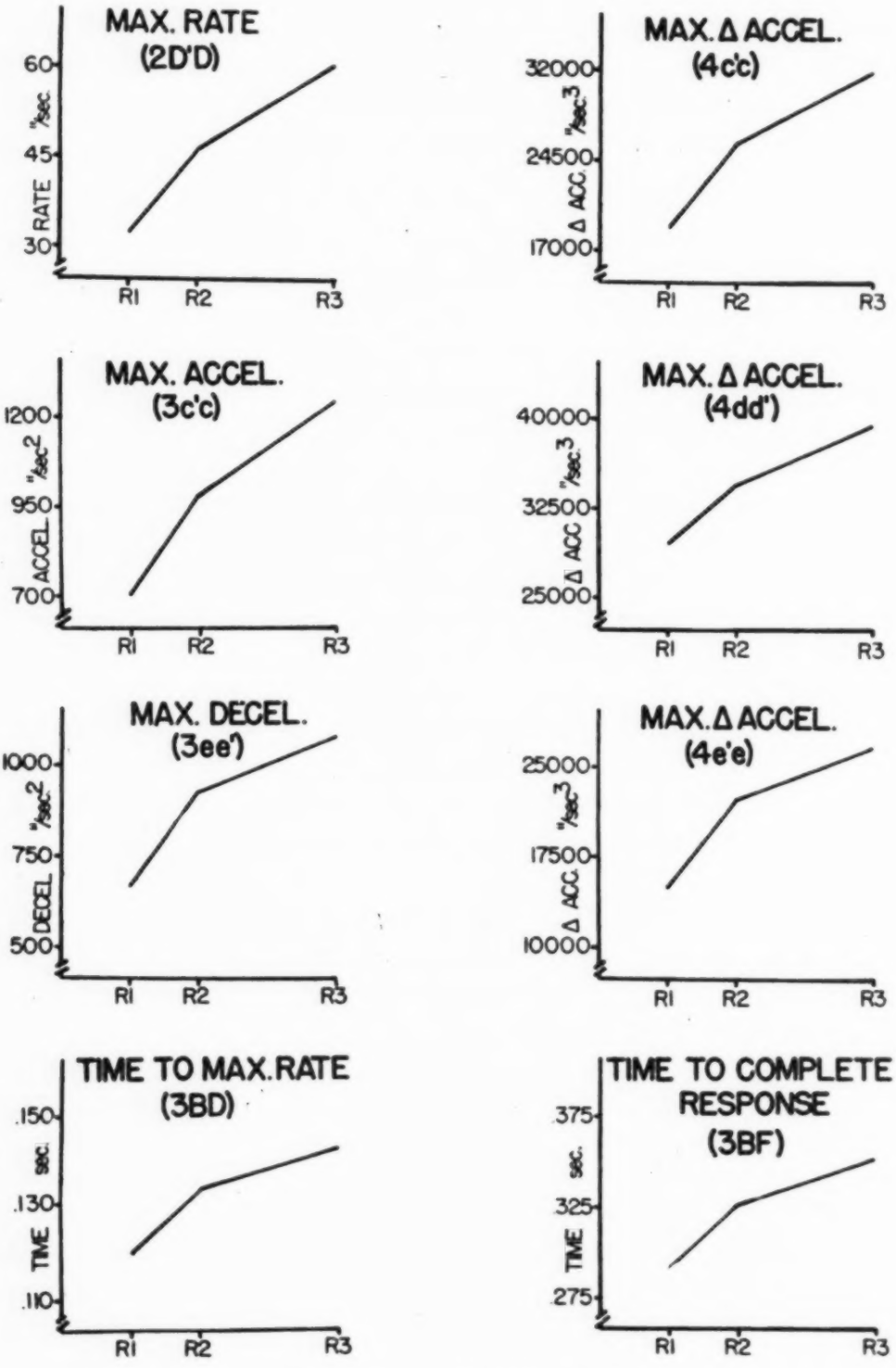


FIG. 3. Averages for all subjects

TABLE I

THE *t*-RATIOS AND LEVELS OF SIGNIFICANCE OF DIFFERENCES BETWEEN NEIGHBORING POINTS ON GRAPHS IN FIG. 3

	Differences between Values for R_1 and R_2		Differences between Values for R_2 and R_3	
	t-ratio	Level of Significance	*t*-ratio	Level of Significance
Max. rate	8.07	below 1 percent	11.90	below 1 percent
Max. acceleration	6.57	below 1 percent	6.00	below 1 percent
Max. deceleration	4.78	below 1 percent	2.60	below 2.8 percent
Δ Acc. (4 c′ c)	4.56	below 1 percent	5.64	below 1 percent
Δ Acc. (4 d d′)	4.17	below 1 percent	2.63	below 2.8 percent
Δ Acc. (4 e′ e)	4.35	below 1 percent	3.21	below 1 percent
Time (3 B D)	7.00	below 1 percent	2.17	below 5.6 percent
Time (3 B F)	3.24	below 1 percent	2.10	below 6.6 percent

time measures, the lower *t*-ratios are the combined result of the small differences between the values, the difficulty of reading the times from the photographs and the high variability of *Ss*.

The ninth measure taken from the photographs was the extent of undershoot or overshoot in the responses to the three sizes of excursions. On the average, the *Ss* overshot R_1 and R_2 by 0.9 in. while they undershot R_3 by 0.2 in. This means that the range of actual excursion lengths was narrower than the range of excursions called for. This, in turn signifies that the shape of the curves obtained would be modified if the data were plotted against the actual displacements instead of the required displacements.

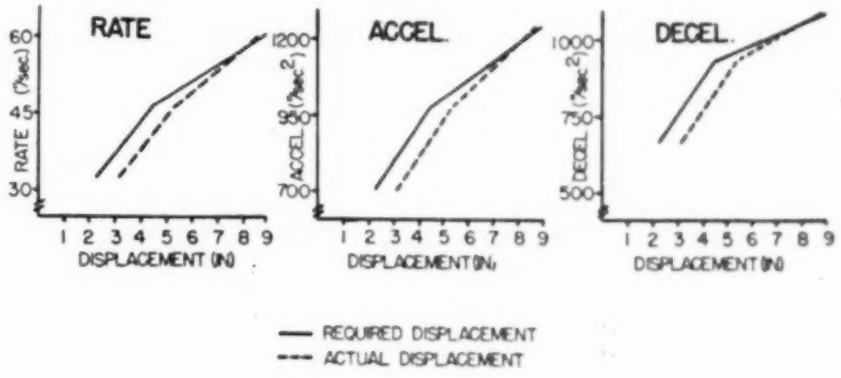


FIG. 4. Maximum rate, maximum acceleration, and maximum deceleration as functions of required and actual hand displacement

The effect upon curve shape may be seen in Fig. 4. Here the values for maximum rate, maximum acceleration and maximum deceleration are plotted against both required hand displacement and actual hand displacement. It is apparent that the curves tend to straighten out when plotted against actual rather than required displacement. The same effect occurs with all of the other measures obtained. This

fact shows clearly that part of the nonlinearity observed in the graphs in Fig. 3 is a function of different response accuracies for target displacements of a different size.

C. Individual differences

Though the response curves for the individual *S*s exhibit the same upward trend with increasing displacement as is found in the average curves, considerable individual variation is to be observed in the slope and shape of the performance curves and in the level of performance. The variation in level of performance is indicated by Table II. Here are shown the averages of the individual *S*s for maximum rate and maximum acceleration achieved with an R_3 excursion.

TABLE II

INDIVIDUAL AVERAGES FOR MAXIMUM RATE AND MAXIMUM ACCELERATION WITH THE R_3 EXCURSION

Subjects	Max. Rate (in./sec.)	Max. Acc. (in./sec.^{2})
Fr.	82.4	1850
Ca.	36.5	612
Bi.	47.7	970
Ag.	40.9	768
Ha.	47.1	676
Jo.	74.6	1570
El.	55.4	864
He.	76.0	2308
Je.	79.5	1610
De.	62.0	1210

It may be noted that the slowest *S* produced a rate of 36.5 in./sec. while the fastest *S* developed a rate of 82.4 in./sec.; a difference of over two times. Maximum acceleration ranged among *S*s from 612 in./sec.^{2} to 2308 in./sec.^{2} This is a difference of almost four times.

DISCUSSION

It might seem useful to try to specify exactly the shapes of the performance curves shown in Fig. 3 in the hope of arriving at some mathematical generalization concerning the rate and acceleration of arm motion in response to the three sizes of displacement. However, as the result of many attempts to fit the curves to formulae, it became apparent that there was no one generalized curve for rate or acceleration or for any of the other measures. The precise forms that the curves assume is a function of the individual *S*s, the particular experimental conditions and the values chosen against which to do the plotting.

It has already been noted that the curves alter their shape when plotted against the values for actual excursion instead of presented displacement. It has also been pointed out that the curve shapes are different for different *S*s. Evidence that the particular experimental situation employed alters the form of the performance curves

is not furnished in this report but is discussed in a previous article (7). Since so many factors enter into the determination of curve shape, mathematical formulation does not appear feasible at present.

However, though no inferences are drawn from the particular nonlinearities observed in this experiment, the universal upward trend of all values with increasing size of excursion does allow certain conclusions. Since the joy stick mechanism employed in this test is nearly frictionless, it is possible to interpret the findings in regard to acceleration and Δ acceleration in terms of force applied by the *S* to the joy stick.

When described in these terms, the results permit the following generalizations: As the target displacement increases, (a) the *S* applies more force in the direction of motion and also more negative or braking force; (b) he applies and removes these forces (positive and negative) at a greater rate; and (c) he applies these forces over a slightly greater period of time.

It is also apparent from the absence of constant or relatively constant rate in the typical acceleration pattern, that from the time the arm starts to move until it stops, the applied force is continuously changing and that it reaches zero only for an instant during the course of motion. This means that the type of movement observed in this experiment is not a 'ballistic' motion where the agonists contract and then relax with force dropping to zero for a period of time before the antagonists come into play (9).

The fact that the movements observed in this study are 'continuously controlled' movements in contrast to 'ballistic' motions, immediately raises the question as to the nature of the control mechanism. It might, at first, be assumed that the sequence of applied forces observed in this experiment arises in response to the changing size of the perceived error as the target is moved toward the hairline. Thus, it might be assumed that the *S* becomes aware of the target displacement, starts to move the joy stick, watches the target move toward the hairline and guides his movement in terms of what he sees.

However, it becomes apparent that this can hardly be the case when the time factor is considered. The average reaction time found in this study was 0.435 sec. Nevertheless, once the arm started to move the force pattern varied continuously. For approximately the first 0.07 sec. force in the direction of motion was applied at an increasing rate, then it was applied at a decreasing rate for another 0.07 sec., then braking force was applied at an increasing rate during the next 0.10 sec., and finally the negative force was applied at a decreasing rate for the last 0.10 sec. of the response. Thus, the arm had passed through all phases of acceleration and had come to rest in a length of time no greater than the reaction time.

Even if it is assumed that the obtained reaction time which is being used as the standard is abnormally long, it is still true that the early phases of acceleration are completed in less time than it would take to perceive and to react selectively to a visual stimulus. This must mean that at least the early phases of the response are not guided continuously by information obtained through vision. The *S* starts to stop before he can react to visual information which tells him how far he has already moved.

It is doubtful if this effect can be accounted for by proprioception or 'muscle feedback.' Though it is conceivable that the reaction time to proprioceptive cues may be shorter than that to visual stimuli, the proprioceptive reaction time would have to be infinitely small to permit kinesthetic control of the continuous variation in applied force found in this study.

Though working with relatively unrefined techniques, Peters and Wenborne (**6**) reached conclusions similar to the above with their stylus study. They suggest that quick arm motions characterized by increasing velocity over the first half and decreasing velocity over the second half are executed without sensory control. Such motions they describe as a 'motor impulse effect.' Craik (**2**) also cites evidence in favor of this type of response.

At the present time, the actual mechanism underlying these responses cannot be described. However, the force patterns that appear suggest the analogy of a program cam control. Using this analogy for what it is and nothing more, we may describe the process of making quick corrections as follows: The *S* perceives an error, selects a cam which is judged on the basis of past experience to best approximate the required force pattern and then starts the mechanism operating. The perception and cam selection require one reaction time. Once started, the cam runs and the continuously varying force pattern which includes stopping forces as well as starting forces appears in proper sequence. Thus, in this analogy, no visual or proprioceptive 'feedback' is required to control the action sequence once a response is underway.

It might further be assumed with considerable logic that what is usually considered continuous tracking is actually a sequence of these patterned 'motor impulse' type responses tied together and successively organized by the visual perception of what has gone before. At any rate our findings suggest that with rapid tracking, control is at least in part, an intermittent rather than a continuous process. The writings of Craik (**2, 3**) and Hick (**5**) have discussed this concept of intermittency in tracking at some length.

At the moment we know next to nothing of the mechanisms which produce and control this intermittency. Research on this problem

is continuing, however, in the hope that eventually we may come to understand the conditions which underlie this aspect of human follow-up behavior.

Summary

1. Visual errors in the two horizontal directions and of three different sizes were produced instantaneously and presented in random order to 10 *Ss*. The errors were corrected through the movement of a joy stick. The joy stick responses were analyzed electrically into time plots of position, rate, acceleration and the third derivative of acceleration (Δ acceleration).

2. Since the control mechanism was nearly frictionless, it is possible to interpret the findings in regard to acceleration and Δ acceleration in terms of force applied by the subject to the joy stick. When interpreted in this way the following generalizations emerge:

a. As the visual error increases, (1) the *S* applies more force in the direction of motion and also more braking force; (2) he applies and removes these forces at a greater rate; and (3) he applies these forces over a slightly greater period of time.

b. Throughout the course of any one motion, force varies continuously.

3. The time relations of these corrective motions are such that it appears that once started, the motions run off without visual or kinesthetic guidance.

4. The latter finding suggests that control in target tracking is an intermittent rather than a continuous process.

(Manuscript received for immediate publication July 15, 1948)

References

1. Barnes, R. M. *Motion and time study.* London: John Wiley & Sons, Inc., 1946.
2. Craik, K. J. W. Theory of the human operator in control systems, Part I. The operator as an engineering system. *Brit. J. Psychol.*, 1947, **38**, 56–61.
3. Craik, K. J. W. Theory of the human operator in control systems, Part II. Man as an element in a control system. *Brit. J. Psychol.*, 1948, **38**, 142–148.
4. de Montpellier, G. Note on acceleration in voluntary hand movements. *Arch. Psychol., Genève*, 1937, **26**, 181–197.
5. Hick, W. E. Discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, 1948, **1**, 36–51.
6. Peters, W., & Wenborne, A. A. The time pattern of voluntary movements. Parts I–III. *Brit. J. Psychol.*, 1936, **26**, 388–406. Parts IV–VI. *Brit. J. Psychol.*, 1936, **27**, 60–73.
7. Searle, L. V., & Taylor, F. V. Studies of tracking behavior. I. Rate and time characteristics of simple corrective movements. *J. exp. Psychol.*, 1948, **38**, 615–631.
8. Slater-Hammel, A. T., & Brown, J. S. Discrete movements in the horizontal plane as a function of their length and direction. *Special Devices Center, Navy Department, Contract N5ori-57 Project 2, Report 2*, October, 1947, 1–24.
9. Stetson, R. H., & McDill, J. A. Mechanism of the different types of movement. *Psychol. Monogr.*, 1923, **32**, 18–40.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The issues listed below are available for sale

Year	Volume	Available Numbers						Price per Number	Price per Volume
1916	1	1	2	3	4	5	6	$1.25	$7.00
1917	2	1	2	3	4	5	6	$1.25	$7.00
1920	3	–	2	3	4	5	6	$1.25	$6.25
1921	4	1	2	3	4	5	6	$1.25	$7.00
1922	5	1	2	3	4	5	6	$1.25	$7.00
1923	6	1	–	3	4	–	–	$1.25	$3.75
1924	7	1	2	–	–	5	–	$1.25	$3.75
1925	8	1	2	3	4	–	6	$1.25	$6.25
1926	9	–	2	3	–	5	6	$1.25	$5.00
1927	10	1	2	3	4	5	6	$1.25	$7.00
1928	11	1	2	3	4	5	6	$1.25	$7.00
1929	12	1	2	3	4	5	6	$1.25	$7.00
1930	13	1	2	3	4	5	6	$1.25	$7.00
1931	14	1	2	3	4	5	6	$1.25	$7.00
1932	15	1	2	3	4	5	6	$1.25	$7.00
1933	16	1	2	3	4	5	6	$1.25	$7.00
1934	17	1	2	3	4	5	6	$1.25	$7.00
1935	18	1	2	3	4	5	6	$1.25	$7.00
1936	19	1	2	3	4	5	6	$1.25	$7.00
1937	20	1	2	3	4	5	6	$1.25	$7.00
	21	1	2	3	4	5	6	$1.25	$7.00
1938	22	1	2	3	4	5	6	$1.25	$7.00
	23	1	2	3	4	5	6	$1.25	$7.00
1939	24	1	2	3	4	5	6	$1.25	$7.00
	25	1	2	3	4	5	6	$1.25	$7.00
1940	26	1	2	3	4	5	6	$1.25	$7.00
	27	1	2	3	4	5	6	$1.25	$7.00
1941	28	1	2	3	4	5	6	$1.25	$7.00
	29	1	2	3	4	5	6	$1.25	$7.00
1942	30	1	2	3	4	5	6	$1.25	$7.00
	31	1	2	3	4	5	6	$1.25	$7.00
1943	32	1	2	3	4	5	6	$1.25	$7.00
	33	1	2	3	4	5	6	$1.25	$7.00
1944	34	1	2	3	4	5	6	$1.25	$7.00
1945	35	1	2	3	4	5	6	$1.25	$7.00
1946	36	1	2	3	4	5	6	$1.25	$7.00
1947	37	1	2	3	4	5	6	$1.25	$7.00
1948	38	By subscription						$1.25	——
								List price, Volumes 1 through 37	$249.00
								30% discount	74.70
								Net price, Volumes 1 through 37	$174.30

Information about the Journal of Experimental Psychology: as noted in the table, the journal was not published in 1918 and 1919; from 1937 to 1943 it was published at the rate of two volumes per year. Only eleven numbers affecting six volumes are out of print in the entire series.

Information about prices: when all numbers are available, the price per volume throughout the series is $7.00. For incomplete volumes, the price is $1.25 for each available number. For foreign postage, $.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of $ 50.00 and over
20% on orders of $100.00 and over
30% on orders of $150.00 and over

Address orders and subscriptions to:

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Avenue, N.W. Washington 5, D. C.

ANNUAL CLUB RATES FOR JOURNALS

American Psychological Association

The American Psychological Association offers its journals in various combinations at special rates. If you are now subscribing to several of the Association's journals at the individual subscription rates, you can secure additional journals or save money by subscribing to one of the combinations described below in the table. At the left are the names of the journals; the second column gives their regular subscription prices; the third column gives their club-rate prices. The reduction in price for each journal is the same no matter in what combination it appears.

The numbers given the clubs are the same as in former years. The prices are approximately the same, except that the *Journal of Consulting Psychology* has been raised in list price from $3.00 to $5.00 per year, and the clubs in which it appears are correspondingly raised in price.

Club 1 includes all ten journals published by the Association. Club 4 includes the seven journals which the Association has published the longest time. The three most recently acquired journals not included in Club 4 are the American Psychologist, the Journal of Comparative and Physiological Psychology (formerly published by Williams & Wilkins), and the Journal of Consulting Psychology (formerly published by the American Association for Applied Psychology). Other combinations can be determined from inspection of the table.

Club rates are available on a January-December basis only.

TABLE OF CLUBS

Journal	List Price	Club Price	Club 1	Club 4	Club 5	Club 6	Club 9
American Psychologist	$7.00	$6.00	x		x		
Journal of Abnormal and Social Psychology	5.00	4.50	x	x	x	x	x
Journal of Applied Psychology	6.00	5.00	x	x	x	x	
Journal of Comparative and Physiological Psychology	7.00	6.00	x				
Journal of Consulting Psychology	5.00	4.50	x		x		
Journal of Experimental Psychology	7.00	6.00	x	x		x	x
Psychological Abstracts	7.00	6.00	x	x	x		
Psychological Bulletin	7.00	6.00	x	x		x	
Psychological Monographs: General and Applied	6.00	5.00	x	x	x	x	x
Psychological Review	5.50	5.00	x	x	x	x	x
List price would be			$62.50	$43.50	$41.50	$36.50	$23.50
Club Price			$54.00	$37.50	$36.00	$31.50	$20.50

AMERICAN PSYCHOLOGICAL ASSOCIATION
1515 Massachusetts Avenue, N. W.
Washington 5, D. C.

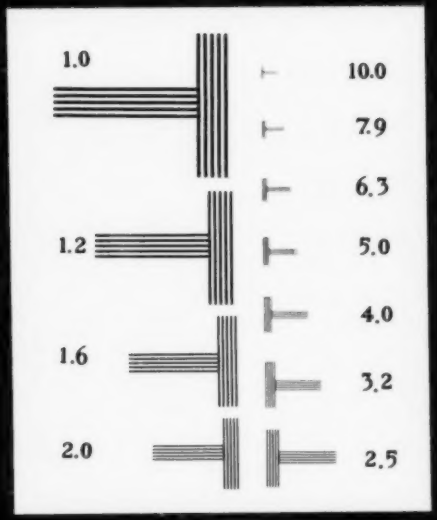
JOURNAL OF EXPERIMENTAL PSYCHOLOGY

VOL. 39, 1949
NOs 1-6

PUB. 276

UNIVERSITY MICROFILMS
ANN ARBOR, MICHIGAN. 1950

RESOLUTION CHART



100 MILLIMETERS

INSTRUCTIONS Resolution is expressed in terms of the lines per millimeter recorded by a particular film under specified conditions. Numerals in chart indicate the number of lines per millimeter in adjacent "T-shaped" groupings.

In microfilming, it is necessary to determine the reduction ratio and multiply the number of lines in the chart by this value to find the number of lines recorded by the film. As an aid in determining the reduction ratio, the line above is 100 millimeters in length. Measuring this line in the film image and dividing the length into 100 gives the reduction ratio. Example: the line is 20 mm. long in the film image, and $100/20 = 5$.

Examine "T-shaped" line groupings in the film with microscope, and note the number adjacent to finest lines recorded sharply and distinctly. Multiply this number by the reduction factor to obtain resolving power in lines per millimeter. Example: 7.9 group of lines is clearly recorded while lines in the 10.0 group are not distinctly separated. Reduction ratio is 5, and $7.9 \times 5 = 39.5$ lines per millimeter recorded satisfactorily. $10.0 \times 5 = 50$ lines per millimeter which are not recorded satisfactorily. Under the particular conditions, maximum resolution is between 39.5 and 50 lines per millimeter.

Resolution, as measured on the film, is a test of the entire photographic system, including lens, exposure, processing, and other factors. These rarely utilize maximum resolution of the film. Vibrations during exposure, lack of critical focus, and exposures yielding very dense negatives are to be avoided.

THIS PUBLICATION IS REPRODUCED BY AGREEMENT WITH THE COPYRIGHT OWNER. EXTENSIVE DUPLICATION OR RESALE WITHOUT PERMISSION IS PROHIBITED.